

84246

**From:** Chan, Christina  
**Sent:** Tuesday, January 14, 2003 3:06 PM  
**To:** Holleran, Anne; STIC-Biotech/ChemLib  
**Subject:** RE: RUSH sequence search for 09/506,079

**Importance:** High

**Please rush. Thanks Chris**

Chris Chan  
TC 1600 New Hire Training Coordinator and SPE 1644  
308-3973  
CM-1, 9B19

*Chris*

-----Original Message-----

**From:** Holleran, Anne  
**Sent:** Tuesday, January 14, 2003 12:12 PM  
**To:** Chan, Christina  
**Subject:** RUSH sequence search for 09/506,079

Please approve and forward to STIC the following RUSH sequence search request. This is for an amendment due this biweek. Thanks.

Please search the following for 09/506,079:

commercial and interference databases oligomer search of SEQ ID NO: 1(aa)

commercial and interference databases oligomer search of SEQ ID NO: 2(aa)

SEQ ID NO: 1 is a fragment of SEQ ID NO: 2.

~~Anne Holleran~~  
AU: 1642  
Tel: 308-8892  
RM: 8e03  
mailbox: 8e12

Point of Contact:  
Barb O'Bryen  
Technical Information Specialist  
STIC CM1 6A05 308-4291

Searcher: *BOB*  
Phone: \_\_\_\_\_  
Location: \_\_\_\_\_  
Date Picked Up: \_\_\_\_\_  
Date Completed: *1-15-03*  
Searcher Prep/Review: \_\_\_\_\_  
Clerical: \_\_\_\_\_  
Online time: \_\_\_\_\_

TYPE OF SEARCH:  
NA Sequences: \_\_\_\_\_  
AA Sequences: \_\_\_\_\_  
Structures: \_\_\_\_\_  
Bibliographic: \_\_\_\_\_  
Litigation: \_\_\_\_\_  
Full text: \_\_\_\_\_  
Patent Family: \_\_\_\_\_  
Other: \_\_\_\_\_

VENDOR/COST (where applic.)  
STN: \_\_\_\_\_  
DIALOG: \_\_\_\_\_  
Questel/Orbit: \_\_\_\_\_  
DRLink: \_\_\_\_\_  
Lexis/Nexis: \_\_\_\_\_  
Sequence Sys.: \_\_\_\_\_  
WWW/Internet: \_\_\_\_\_  
Other (specify): \_\_\_\_\_

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

### OM protein - protein search, using sw model

Run on: January 14, 2003, 17:11:49 ; Search time 5.0763 seconds

(without alignments)  
457.894 Million cell updates/sec

Title: US-09-506-079E-1  
Perfect score: 79

Sequence: 1 GKHSSXPPRPAAVPVPAQRXQP . . . . . VGRGXDPDAHVAVLRSYEG 79

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 262574 seqs, 29422922 residues

Word size : 0

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

### Post-processing: Listing first 45 summaries

Database : Issued\_Patents\_AA:\*

1: /cgn2\_6/podata/1/iaa/5A.COMB.pep:\*

2: /cgn2\_6/podata/1/iaa/5B.COMB.pep:\*

3: /cgn2\_6/podata/1/iaa/6A.COMB.pep:\*

4: /cgn2\_6/podata/1/iaa/6B.COMB.pep:\*

5: /cgn2\_6/podata/1/iaa/6C.PETUS.COMB.pep:\*

6: /cgn2\_6/podata/1/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

### SUMMARIES

Result No.	Score	Query	Length	DB	ID	Description
1	17	21.5	79	4	US-09-630-155-1	Sequence 1, Appli
2	17	21.5	419	4	US-09-630-155-2	Sequence 2, Appli
3	7	8.9	635	2	US-09-014-969-11	Sequence 11, Appli
4	6	7.6	12	2	US-08-492-146	Sequence 146, Appli
5	6	7.6	15	1	US-08-036-555B-49	Sequence 49, Appli
6	6	7.6	15	1	US-08-469-569-49	Sequence 49, Appli
7	6	7.6	15	1	US-08-249-322A-49	Sequence 49, Appli
8	6	7.6	15	1	US-08-469-526A-49	Sequence 49, Appli
9	6	7.6	15	2	US-08-734-591A-49	Sequence 49, Appli
10	6	7.6	15	2	US-08-469-660-49	Sequence 49, Appli
11	6	7.6	15	3	US-08-341-018-82	Sequence 82, Appli
12	6	7.6	15	4	US-08-470-335-49	Sequence 49, Appli
13	6	7.6	15	4	US-08-735-021-49	Sequence 49, Appli
14	6	7.6	15	4	US-08-734-664A-49	Sequence 49, Appli
15	6	7.6	15	4	US-08-470-339-49	Sequence 49, Appli
16	6	7.6	15	4	US-08-467-602-49	Sequence 49, Appli
17	6	7.6	15	5	PCT-US94-05083C-46	Sequence 46, Appli
18	6	7.6	15	5	PCT-US95-00846A-49	Sequence 49, Appli
19	6	7.6	16	1	US-08-469-569-37	Sequence 37, Appli
20	6	7.6	16	1	US-08-469-569-37	Sequence 37, Appli
21	6	7.6	16	1	US-08-249-322A-37	Sequence 37, Appli
22	6	7.6	16	1	US-08-469-526A-37	Sequence 37, Appli
23	6	7.6	16	2	US-08-734-591A-37	Sequence 37, Appli
24	6	7.6	16	2	US-08-469-660-37	Sequence 37, Appli
25	6	7.6	16	4	US-08-470-335-49	Sequence 37, Appli
26	6	7.6	16	4	US-08-735-021-37	Sequence 37, Appli
27	6	7.6	16	4	US-08-734-664A-37	Sequence 37, Appli

### ALIGNMENTS

RESULT 1  
US-09-630-155-1  
Sequence 1, Application US/09630155  
; Patent No. 6414130  
GENERAL INFORMATION:  
APPLICANT: Doherty, Joni Kristin and Gail M. Clinton  
TITLE OF INVENTION: HER-2 BINDING ANTAGONISTS  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:  
NAME: Davison, Barry L.  
REGISTRATION NUMBER: 47,309  
REFERENCE/DOCKET NUMBER: 49321-10  
TELECOMMUNICATION INFORMATION:  
STREET: 1501 Fourth Avenue, 2600 Century Square  
CITY: Seattle  
STATE: Washington  
COUNTRY: U.S.A.  
ZIP: 98101  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/630,155  
FILING DATE: 15-Jan-2001  
CLASSIFICATION: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Davison, Barry L.  
REGISTRATION NUMBER: 47,309  
REFERENCE/DOCKET NUMBER: 49321-10  
TELEPHONE: 206 628-7621  
TELEFAX: 206 628-7699  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 79  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: unknown  
MOLECULE TYPE: HER-2 ECD antagonist  
SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
US-09-630-155-1  
Query Match Best Local Similarity 21.5%; Score 17; DB 4; Length 79;  
Matches 17; Conservative 100.0%; Pred. No. 5e-10; Indels 0; Gaps 0;

RESULT 2  
US-09-630-155-2  
Sequence 2, Application US/09630155  
; Patient No. 6414130  
; GENERAL INFORMATION:  
; APPLICANT: Doherty, Joni Kristin and Gail M. Clinton  
; TITLE OF INVENTION: HER-2 BINDING ANTAGONISTS  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: DAVIS WRIGHT TREMAYNE LLP  
; STREET: 1501 Fourth Avenue, 2600 Century Square  
; CITY: Seattle  
; STATE: Washington  
; COUNTRY: U.S.A.  
; ZIP: 98101  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/014,969  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Spuriger, Suzanne A.  
; REGISTRATION NUMBER: 411323  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 498-8284  
; TELEFAX: (617) 876-5851  
; INFORMATION FOR SEQ ID NO: 11:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 635 amino acids  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; TYPE: amino acid  
; MOLECULE TYPE: protein  
; US-09-014-969-11  
; INFORMATION FOR SEQ ID NO: 2:  
; LENGTH: 419  
; STRANDEDNESS: single  
; TOPOLOGY: unknown  
; MOLECULE TYPE: polypeptide  
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
; US-09-630-155-2  
Query Match 21.5%; Score 17; DB 4; Length 419;  
Best Local Similarity 100.0%; Pred. No. 2.2e-09; Indels 0; Gaps 0;  
Matches 17; Conservative 0; Mismatches 0;  
Qy 37 VSAFYSLPLAPSPTV 53  
Db 377 VSAFYSLPLAPSPTV 393  
RESULT 3  
US-09-014-969-11  
; Sequence 11, Application US/09014969  
; General Information:  
; APPLICANT: Jacobs, Kenneth  
; APPLICANT: McCoy, John M.  
; APPLICANT: Lavallie, Edward R.  
; APPLICANT: Racine, Lisa A.  
; APPLICANT: Merberg, David  
; APPLICANT: Tracy, Maurice  
; APPLICANT: Spaulding, Vicki  
; APPLICANT: Agostino, Michael J.  
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES  
; NUMBER OF SEQUENCES: 32  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genetics Institute, Inc.  
; STREET: 87 CambridgePark Drive  
; CITY: Cambridge  
; STATE: MA  
; COUNTRY: U.S.A.  
; ZIP: 02140  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC\ DOS\ MS\ DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/811,492  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIORITY APPLICATION DATA:  
; APPLICATION NUMBER: US 08/496,247  
; PRIORITY APPLICATION DATA:

APPLICATION NUMBER: 07/1963-703	
FILING DATE: 28-JUN-1995	
CLASSIFICATION: 435	
PRIORITY APPLICATION DATA:	
APPLICATION NUMBER: US 08/146, 885	
FILING DATE: 03-NOV-1993	
CLASSIFICATION: 435	
PRIORITY APPLICATION DATA:	
APPLICATION NUMBER: US 08/004, 139	
FILING DATE: 09-DEC-1992	
CLASSIFICATION: 435	
ATTORNEY/AGENT INFORMATION:	
NAME: Williams, Gregory D	
REGISTRATION NUMBER: 30901	
REFERENCE/DOCKET NUMBER: NEB-036CA	
TELECOMMUNICATION INFORMATION:	
TELEPHONE: 508-927-5054	
TELEFAX: 509-927-1705	
TELEX:	
INFORMATION FOR SEQ ID NO: 146:	
SEQUENCE CHARACTERISTICS:	
LENGTH: 12 amino acids	
TYPE: amino acid	
STRANDEDNESS: single	
TOPOLOGY: linear	
MOLECULE TYPE: peptide	
S-08-811-492-146	
Query Match 7.6%; Score 6; DB 2; Length 12;	
Best Local Similarity 100.0%; Pred. No. 8.1;	Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Y 45 LAPLSP 50	
7 LAPLSP 12	
RESULT 5	
Sequence 49, Application US/08036555B	
Patient No. 5530109	
GENERAL INFORMATION:	
APPLICANT: Goodearl, Andrew; Stroobant, Paul;	
APPLICANT: Minghetti, Luisa; Waterfield, Michael; Marchioni, Mark;	
APPLICANT: Chen, Mai; Su, Hiles, Ian	
TITLE OF INVENTION: Glial Mitogenic Factors, Their Preparation and Use	
TITLE OF INVENTION: Preparation and Use	
NUMBER OF SEQUENCES: 184	
CORRESPONDENCE ADDRESS:	
ADDRESSEE: Felfe & Lynch	
STREET: 805 Third Avenue	
CITY: New York City	
STATE: New York	
COUNTRY: USA	
ZIP: 10022	
COMPUTER READABLE FORM:	
COMPUTER: IBM	
OPERATING SYSTEM: PC-DOS	
SOFTWARE: Wordperfect	
CURRENT APPLICATION DATA:	
APPLICATION NUMBER: US/08/036, 555B	
FILING DATE: 24-MAR-1993	
CLASSIFICATION: 435	
PRIORITY APPLICATION DATA:	
APPLICATION NUMBER: -07/965, 173	
FILING DATE: 23-OCT-1992	
PRIOR APPLICATION DATA:	
APPLICATION NUMBER: 07/940, 389	
FILING DATE: 03-SEP-1992	
PRIOR APPLICATION DATA:	
APPLICATION NUMBER: 07/907, 138	
FILING DATE: 30-JUN-1992	
PRIOR APPLICATION DATA:	
APPLICATION NUMBER: 07/863, 703	
FILING DATE: 03-APRIL-1992	
PRIOR APPLICATION DATA:	
APPLICATION NUMBER: U K. 91 07566.3	
FILING DATE: 10-APRIL-1991	
PRIMER APPLICATION DATA:	
APPLICATION NUMBER: 07/1963-703	
FILING DATE: 03-APRIL-1992	
PRIOR APPLICATION DATA:	
APPLICATION NUMBER: U K. 91 07566.3	
FILING DATE: 10-APRIL-1991	
ATTORNEY/AGENT INFORMATION:	
NAME: Tsai, Christine H.	
REGISTRATION NUMBER: 34, 266	
REFERENCE/DOCKET NUMBER: LUD 5250.4	
TELECOMMUNICATION INFORMATION:	
TELEPHONE: (212) 688-9200	
TELEFAX: (212) 838-3884	
INFORMATION FOR SEQ ID NO: 49:	
SEQUENCE CHARACTERISTICS:	
LENGTH: 15	
TYPE: amino acid	
STRANDEDNESS:	
TOPOLOGY: linear	
US-08-036-555B-49	
Query Match 7.6%; Score 6; DB 1; Length 15;	
Best Local Similarity 100.0%; Pred. No. 10;	Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 56 SPVSVG 61	
Db 2 SPVSVG 7	
RESULT 6	
Sequence 49, Application US/08469569	
Patient No. 5406032	
GENERAL INFORMATION:	
APPLICANT: Goodearl, Andrew; Stroobant, Paul;	
APPLICANT: Minghetti, Luisa; Waterfield, Michael; Marchioni, Mark;	
APPLICANT: Chen, Mai; Su, Hiles, Ian	
TITLE OF INVENTION: Glial Mitogenic Factors, Their Preparation and Use	
TITLE OF INVENTION: Preparation and Use	
NUMBER OF SEQUENCES: 184	
CORRESPONDENCE ADDRESS:	
ADDRESSEE: Felfe & Lynch	
STREET: 805 Third Avenue	
CITY: New York City	
STATE: New York	
COUNTRY: USA	
ZIP: 10022	
COMPUTER READABLE FORM:	
COMPUTER: IBM	
OPERATING SYSTEM: PC-DOS	
SOFTWARE: Wordperfect	
CURRENT APPLICATION DATA:	
APPLICATION NUMBER: US/08/469, 569	
FILING DATE: 06-JUN-1995	
CLASSIFICATION: 530	
PRIOR APPLICATION DATA:	
APPLICATION NUMBER: 08/036, 555	
FILING DATE: 24-MAR-1993	
APPLICATION NUMBER: 07/965, 173	
FILING DATE: 23-OCT-1992	
PRIOR APPLICATION DATA:	
APPLICATION NUMBER: 07/940, 389	
FILING DATE: 03-SEP-1992	
PRIOR APPLICATION DATA:	
APPLICATION NUMBER: 07/907, 138	
FILING DATE: 30-JUN-1992	
PRIOR APPLICATION DATA:	
APPLICATION NUMBER: 07/863, 703	
FILING DATE: 03-APRIL-1992	
PRIOR APPLICATION DATA:	
APPLICATION NUMBER: U K. 91 07566.3	
FILING DATE: 10-APRIL-1991	

## ATTORNEY / AGENT INFORMATION:

NAME: Tsai, Christine H.

REGISTRATION NUMBER: 34,266

REFERENCE/DOCKET NUMBER: LUD 5250.4

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 688-9200

TELEFAX: (212) 838-3884

INFORMATION FOR SEQ ID NO: 49:

SEQUENCE CHARACTERISTICS:

LENGTH: 15

TYPE: amino acid

STRANDEDNESS: Linear

TOPOLOGY: Linear

US-08-469-569-49

Query Match 7.6%; Score 6; DB 1; Length 15;

Best Local Similarity 100.0%; Pred. No. 10;

Matches 6; Conservative 0; Mismatches 0;

Indels 0; Gaps 0;

QY 56 SPVSVG 61

Db 2 SPVSVG 7

US-08-249-322A-49

Query Match 7.6%; Score 6; DB 1; Length 15;

Best Local Similarity 100.0%; Pred. No. 10;

Matches 6; Conservative 0; Mismatches 0;

Indels 0; Gaps 0;

QY 56 SPVSVG 61

Db 2 SPVSVG 7

RESULT 7

US-08-249-322A-49

Sequence 49, Application US/08249322A

PATENT NO. 5710930

GENERAL INFORMATION:

APPLICANT: Goodearl, Andrew; Stroobant, Paul;

APPLICANT: Minchetti, Luisa; Waterfield, Michael; Marchionni, Mark;

APPLICANT: Chen, Maiq Su; Hiles, Ian

APPLICANT: Glial Mitogenic Factors, Their

TITLE OF INVENTION: Glial Mitogenic Factors, Their

TITLE OF INVENTION: Preparation and Use

NUMBER OF SEQUENCES: 184

CORRESPONDENCE ADDRESS:

ADDRESSEE: Relfe &amp; Lynch

STREET: 805 Third Avenue

CITY: New York City

STATE: New York

COUNTRY: USA

ZIP: 10022

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM

OPERATING SYSTEM: PC-DOS

SOFTWARE: Wordperfect

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/249,322A

FILING DATE: 26-MAY-1994

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/036,555

FILING DATE: 24-MAR-1993

FILING DATE: 24-MAR-1993

FILING DATE: 23-OCT-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/940,389

FILING DATE: 03-SEP-1992

FILING DATE: 03-APRIL-1992

FILING DATE: 07/907,138

FILING DATE: 30-JUN-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/863,703

FILING DATE: 10-APR-1991

ATTORNEY / AGENT INFORMATION:

NAME: Tsai, Christine H.

REGISTRATION NUMBER: 34,266

REFERENCE/DOCKET NUMBER: LUD 250.4

## TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 688-9200

TELEFAX: (212) 838-3884

INFORMATION FOR SEQ ID NO: 49:

SEQUENCE CHARACTERISTICS:

LENGTH: 15

TYPE: amino acid

STRANDEDNESS: Linear

TOPOLOGY: Linear

US-08-249-322A-49

Query Match 7.6%; Score 6; DB 1; Length 15;

Best Local Similarity 100.0%; Pred. No. 10;

Matches 6; Conservative 0; Mismatches 0;

Indels 0; Gaps 0;

QY 56 SPVSVG 61

Db 2 SPVSVG 7

RESULT 8

US-08-469-526A-49

Sequence 49, Application US/08469526A

PATENT NO. 5798849

GENERAL INFORMATION:

APPLICANT: Goodearl, Andrew

APPLICANT: Stroobant, Paul

APPLICANT: Minchetti, Luisa

APPLICANT: Waterfield, Michael

APPLICANT: Marchionni, Mark

APPLICANT: Chen, Maiq Su

APPLICANT: Hiles, Ian

APPLICANT: Glial Mitogenic Factors, Their

TITLE OF INVENTION: Glial Mitogenic Factors, Their

TITLE OF INVENTION: PREPARATION AND USE

NUMBER OF SEQUENCES: 187

CORRESPONDENCE ADDRESS:

ADDRESSEE: Clark &amp; Elbing LLP

STREET: 176 Federal Street

CITY: Boston

STATE: MA

COUNTRY: USA

ZIP: 02110

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FASISEQ for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/469,526A

FILING DATE: 06-JUNE-1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/036,555

FILING DATE: 24-MAR-1993

FILING DATE: 24-MAR-1993

FILING DATE: 23-OCT-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/940,389

FILING DATE: 03-SEP-1992

FILING DATE: 03-APRIL-1992

FILING DATE: 07/907,138

FILING DATE: 30-JUN-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/863,703

FILING DATE: 10-APR-1991

ATTORNEY / AGENT INFORMATION:

NAME: Baker-Brady, Kristina

REGISTRATION NUMBER: 39,109

REFERENCE/DOCKET NUMBER: 04585/00200A

TELECOMMUNICATION INFORMATION:

TELEPHONE: 617-428-0200

TELEFAX: 617-428-7045

INFORMATION FOR SEQ ID NO: 49:

SEQUENCE CHARACTERISTICS:

RESULT 9  
US-08-734-591A-49  
; Sequence 49, Application US/08734591A  
; Patient No. 5854220  
GENERAL INFORMATION:  
APPLICANT: Goodearl, Andrew  
APPLICANT: Stroobant, Paul  
APPLICANT: Minghetti, Luisa  
APPLICANT: Waterfield, Michael  
APPLICANT: Hiles, Ian  
APPLICANT: Marchionni, Mark  
APPLICANT: Chen, Mario  
TITLE OF INVENTION: GLIAL MITOGENIC FACTORS, THEIR PREPARATION AND USE  
NUMBER OF SEQUENCES: 187  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Clark & Elbing LLP  
STREET: 176 Federal Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: U.S.A.  
ZIP: 02110

COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" DISKETTE, 1.44 MB  
COMPUTER: IBM Compatible Pentium  
OPERATING SYSTEM: Windows95  
SOFTWARE: Wordperfect (Version 7.0)

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/734,591A  
FILING DATE: 22-OCT-1996  
CLASSIFICATION: 536

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/470,335  
FILING DATE: 08-JUN-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/036,555  
FILING DATE: 03-MAR-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/955,173  
FILING DATE: 23-OCT-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/940,389  
FILING DATE: 03-SEP-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/907,138  
FILING DATE: 30-JUN-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/863,703  
FILING DATE: 03-APR-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: UK 91 07566.3  
FILING DATE: 10-APR-1991

ATTORNEY/AGENT INFORMATION:  
NAME: Bicker-Brady, Kristina  
REGISTRATION NUMBER: 39,09  
REFERENCE/DOCKET NUMBER: 04585/00200P  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 428-0200  
TELEFAX: 200154  
INFORMATION FOR SEQ ID NO: 49:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear

Query Match 7.6%; Score 6; DB 1; Length 15;  
Best Local Similarity 100.0%; Pred. No. 10;  
Matches 6; Conservative 0; Mismatches 0;  
Indels 0; Gaps 0;

RESULT 10  
US-08-469-660-49  
; Sequence 49, Application US/08469660  
; Patent No. 5876973  
GENERAL INFORMATION:  
APPLICANT: Gwynne, David I.; Marchionni, Mark;  
APPLICANT: McGuire, Robert N.  
TITLE OF INVENTION: INHIBITORS OF CELL PROLIFERATION,  
TITLE OF INVENTION: THEIR PREPARATION AND USE  
NUMBER OF SEQUENCES: 184  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: Massachusetts  
ZIP: 0211-2804

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage  
COMPUTER: IBM  
OPERATING SYSTEM: PC-DOS  
SOFTWARE: Wordperfect

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/469,660  
FILING DATE:  
CLASSIFICATION: 435

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/011,396  
FILING DATE: 20-JAN-1993

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/984,085  
FILING DATE: 01-DEC-1992

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/951,747  
FILING DATE: 25-SEP-1992

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/927,337  
FILING DATE: 10-AUG-1992

ATTORNEY/AGENT INFORMATION:  
NAME: Clark, Paul T.  
REGISTRATION NUMBER: 30,162  
REFERENCE/DOCKET NUMBER: 04585/017004  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 542-5070  
TELEFAX: 200154  
INFORMATION FOR SEQ ID NO: 49:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear

Query Match 7.6%; Score 6; DB 2; Length 15;  
Best Local Similarity 100.0%; Pred. No. 10;

Qy	56	SPVSVG	61	Db	2	SPVSVG	7	Db	2	SPVSVG	7
Qy				Qy	56	SPVSVG	61	Qy	56	SPVSVG	61
Db	2	SPVSVG	7	Db	2	SPVSVG	7	Db	2	SPVSVG	7
RESULT 11											
US-08-341-018-82											
Sequence 49, Application US/08341018A											
; Sequence 49, Application US/08341018A											
; Patent No. 6194377											
; GENERAL INFORMATION:											
; APPLICANT: Gwynne, David I.											
; APPLICANT: Mananthappa, Nagesh K.											
; APPLICANT: Marchionni, Mark A.											
; APPLICANT: Berningham-McDonogh, Olivia											
; APPLICANT: Goldin, Stanley M.											
; APPLICANT: McBurney, Robert N.											
; TITLE OF INVENTION: USE OF NEUREGULINS AS MODULATORS OF											
; TITLE OF INVENTION: CELLULAR COMMUNICATION											
; FILE REFERENCE: 04585/041001											
; CURRENT APPLICATION NUMBER: US/08/341,018A											
; CURRENT FILING DATE: 1994-11-17											
; NUMBER OF SEQ ID NOS: 87											
; SOFTWARE: FastSEQ for Windows Version 4.0											
; SEQ ID NO 82											
; LENGTH: 15											
Oy	56	SPVSVG	61	Oy	56	SPVSVG	61	Oy	56	SPVSVG	61
Db	2	SPVSVG	7	Db	2	SPVSVG	7	Db	2	SPVSVG	7
RESULT 12											
US-08-470-335-49											
Sequence 49, Application US/08470335F											
; Sequence 49, Application US/08470335F											
; Patent No. 6147190											
; GENERAL INFORMATION:											
; APPLICANT: Goodearl, Andrew											
; APPLICANT: Stroobant, Paul											
; APPLICANT: Minghetti, Luisa											
; APPLICANT: Waterfield, Michael											
; APPLICANT: Marchionni, Mark											
; APPLICANT: Hiles, Ian											
; APPLICANT: Chen, Mario S.											
; APPLICANT: Miles, Ian											
; TITLE OF INVENTION: GLIAL MITOGENIC FACTORS, THEIR											
; PREPARATION AND USE											
; FILE REFERENCE: 04585/00200B											
; CURRENT APPLICATION NUMBER: US/08/470,335F											
; CURRENT FILING DATE: 1995-06-06											
; EARLIER APPLICATION NUMBER: 08/036,555											
; EARLIER FILING DATE: 1993-03-24											
; EARLIER APPLICATION NUMBER: 07/863,703											
; EARLIER FILING DATE: 1992-10-23											
; EARLIER APPLICATION NUMBER: 07/940,389											
; EARLIER FILING DATE: 1992-09-03											
; EARLIER APPLICATION NUMBER: 07/907,138											
; EARLIER FILING DATE: 1992-06-30											
; EARLIER APPLICATION NUMBER: 07/863,703											
; EARLIER FILING DATE: 1992-04-03											
; NUMBER OF SEQ ID NOS: 192											
; SOFTWARE: FastSEQ for Windows Version 3.0											
; SEQ ID NO 49											
; LENGTH: 15											
Oy	56	SPVSVG	61	Oy	56	SPVSVG	61	Oy	56	SPVSVG	61
Db	2	SPVSVG	7	Db	2	SPVSVG	7	Db	2	SPVSVG	7
RESULT 13											
US-08-735-021-49											
Sequence 49, Application US/08735021B											
; Sequence 49, Application US/08735021B											
; Patent No. 6194377											
; GENERAL INFORMATION:											
; APPLICANT: Goodearl, Andrew											
; APPLICANT: Stroobant, Paul											
; APPLICANT: Minghetti, Luisa											
; APPLICANT: Waterfield, Michael											
; APPLICANT: Marchionni, Mark											
; APPLICANT: Hiles, Ian											
; APPLICANT: Chen, Mario S.											
; APPLICANT: Miles, Ian											
; TITLE OF INVENTION: GLIAL MITOGENIC FACTORS, THEIR											
; PREPARATION AND USE											
; FILE REFERENCE: 04585/00200B											
; CURRENT APPLICATION NUMBER: US/08/470,335F											
; CURRENT FILING DATE: 1995-06-06											
; EARLIER APPLICATION NUMBER: 08/036,555											
; EARLIER FILING DATE: 1993-03-24											
; NUMBER OF SEQ ID NOS: 252											
; SOFTWARE: FastSEQ for Windows Version 4.0											
; SEQ ID NO 49											
; LENGTH: 15											
Oy	56	SPVSVG	61	Oy	56	SPVSVG	61	Oy	56	SPVSVG	61
Db	2	SPVSVG	7	Db	2	SPVSVG	7	Db	2	SPVSVG	7
RESULT 14											
US-08-734-664A-49											
Sequence 49, Application US/08734664A											
; Sequence 49, Application US/08734664A											
; Patent No. 6204241											
; GENERAL INFORMATION:											
; APPLICANT: Goodearl, Andrew											
; APPLICANT: Stroobant, Paul											
; APPLICANT: Minghetti, Luisa											
; APPLICANT: Waterfield, Michael											
; APPLICANT: Marchionni, Mark											
; APPLICANT: Chen, Mario S.											
; APPLICANT: Miles, Ian											
; TITLE OF INVENTION: GLIAL MITOGENIC FACTORS, THEIR											
; PREPARATION AND USE											
; FILE REFERENCE: 04585/00200B											
; CURRENT APPLICATION NUMBER: US/08/470,335F											
; CURRENT FILING DATE: 1995-06-06											
; EARLIER APPLICATION NUMBER: 08/036,555											
; EARLIER FILING DATE: 1993-03-24											
; NUMBER OF SEQ ID NOS: 187											
; SOFTWARE: FastSEQ for Windows Version 4.0											
; SEQ ID NO 49											
; LENGTH: 15											
Oy	56	SPVSVG	61	Oy	56	SPVSVG	61	Oy	56	SPVSVG	61
Db	2	SPVSVG	7	Db	2	SPVSVG	7	Db	2	SPVSVG	7
RESULT 15											
US-08-470-335-49											
Query Match 7.6%; Score 6; DB 4; Length 15;											
Best Local Similarity 100.0%; Pred. No. 10;											
Matches 6; Conservative 0; Mismatches 0;											
Indels 0; Gaps 0;											
ADDRESSEE: Clark & Elbing LLP											
STREET: 176 Federal Street											
CITY: Boston											
STATE: Massachusetts											
COUNTRY: U.S.A.											
ZIP: 02110											
ZIPPER READABLE FORM:											
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb											

COMPUTER: IBM Compatible Pentium  
 OPERATING SYSTEM: Windows 95  
 SOFTWARE: FastSeq Version 2.0  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/734,664A  
 FILING DATE: 22-OCT-1996  
 CLASSIFICATION: 536  
 PRIORITY APPLICATION DATA:  
 APPLICATION NUMBER: 08/249,322  
 FILING DATE: 26-MAY-1994  
 PRIORITY APPLICATION DATA:  
 APPLICATION NUMBER: 08/036,555  
 FILING DATE: 24-MAR-1993  
 PRIORITY APPLICATION DATA:  
 APPLICATION NUMBER: 07/965,173  
 FILING DATE: 23-OCT-1992  
 PRIORITY APPLICATION DATA:  
 APPLICATION NUMBER: 07/940,389  
 FILING DATE: 03-SEP-1992  
 PRIORITY APPLICATION DATA:  
 APPLICATION NUMBER: 07/907,138  
 FILING DATE: 30-JUN-1992  
 PRIORITY APPLICATION DATA:  
 APPLICATION NUMBER: 07/863,703  
 FILING DATE: 03-APR-1992  
 PRIORITY APPLICATION DATA:  
 APPLICATION NUMBER: UK 91 07566.3  
 FILING DATE: 10-APR-1991  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Bieker-Brady, Kristina  
 REGISTRATION NUMBER: 39,109  
 REFERENCE/DOCKET NUMBER: 04585/002000J  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (617) 428-0200  
 TELEFAX: (617) 428-7045  
 TELEX:  
 TOPOLOGY: linear  
 US-08-734-664A-49

RESULT 15  
 US-08-470-339-49  
 Sequence 49, Application us/08470339C  
 Patent No. 632206  
 GENERAL INFORMATION:  
 APPLICANT: GOODRILL, ANDREW  
 APPLICANT: STROOBANT, PAUL  
 APPLICANT: MINGETTI, LUISA  
 APPLICANT: WATERFIELD, MICHAEL  
 APPLICANT: MARCTIONI, MARK  
 APPLICANT: CHEN, MARIO S.  
 APPLICANT: HILES, IAN  
 TITLE OF INVENTION: GLIAL MITOGENIC FACTORS, THEIR  
 TITLE OF INVENTION: PREPARATION AND USE  
 FILE REFERENCE: 04585/002008  
 CURRENT APPLICATION NUMBER: US/08/470,339C  
 CURRENT FILING DATE: 1995-06-06  
 EARLIER APPLICATION NUMBER: 08/036,555  
 EARLIER FILING DATE: 1993-03-24  
 EARLIER APPLICATION NUMBER: 07/940,389

---

EARLIER FILING DATE: 1992-09-03  
 EARLIER APPLICATION NUMBER: 07/907,138  
 EARLIER FILING DATE: 1992-06-30  
 EARLIER APPLICATION NUMBER: 07/863,703  
 EARLIER FILING DATE: 1992-04-03  
 EARLIER APPLICATION NUMBER: 91 07566.3 GB  
 EARLIER FILING DATE: 1999-04-10  
 NUMBER OF SEQ ID NOS: 226  
 SOFTWARE: FastSEQ for Windows Version 4.0  
 SEQ ID NO 49  
 LENGTH: 15  
 TYPE: PRT  
 ORGANISM: Bos taurus  
 US-08-470-339-49

Query Match	Score	DB	Length	Length	Indels	Gaps	0;
QY	56	SPVSVG	61	11111	0;	0;	
Db	2	SPVSVG	7				

Search completed: January 14, 2003, 17:15:45  
 Job time : 6.0763 secs

**THIS PAGE BLANK (USPTO)**

Gencore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

## OM protein - protein search, using sw model1

Run on: January 14, 2003, 17:10:43 ; Search time 46.2751 Seconds

(without alignments)  
1865.663 Million cell updates/secTitle: US-09-506-079E-2  
Perfect score: 419  
Sequence: 1 MELAALCRWGLLALLPPGA.....VGRGXDPDAIKAVAVLSRVEG 419

Scoring table: OLIGO Gapop 60.0 , Gapext 60.0

Searched: 671580 seqs, 206047115 residues

Word size : 0

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : SPTRREMBL\_21:  
1: sp\_archaea:  
2: sp\_bacteria:  
3: sp\_fungi:  
4: sp\_human:  
5: sp\_invertebrate:  
6: sp\_mammal:  
7: sp\_mtc:  
8: sp\_organelle:  
9: sp\_phage:  
10: sp\_plant:  
11: sp\_rabbit:  
12: sp\_virus:  
13: sp\_vertebrate:  
14: sp\_unclassified:  
15: sp\_virus:  
16: sp\_bacteria:  
17: sp\_archeap:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match Length DB ID	Description
1	83	19.8 419 4 Q9UK79	Q9UK79 homo sapien
2	51	12.2 149 6 Q9BG66	Q9BG66 oryctoferus
3	50	11.9 1259 6 Q18735	Q18735 canis familiaris
4	9	2.1 478 11 Q9ESE0	Q9ESE0 rattus norvegicus
5	9	2.1 527 13 Q9PS36	Q9PS36 gallus gallus
6	9	2.1 599 13 Q9PSH2	Q9PSH2 gallus gallus
7	9	2.1 643 11 Q9ERV5	Q9ERV5 mus musculus
8	9	2.1 655 11 Q9WVF5	Q9WVF5 mus musculus
9	2.1 1209 11 Q9X70	Q9X70 rattus norvegicus	
10	9	2.1 1210 11 Q9EP98	Q9EP98 mus musculus
11	8	1.9 48 11 Q63365	Q63365 rattus norvegicus
12	8	1.9 63 11 Q9ZI6	Q9ZI6 rattus norvegicus
13	8	1.9 63 11 Q9QZ15	Q9QZ15 rattus norvegicus
14	8	1.9 277 11 Q08558	Q08558 rattus norvegicus
15	8	1.9 277 11 Q924V3	Q924V3 cricetulus
16	8	280 13 Q8UVJ2	Q8UVJ2 oreochromis

## ALIGNMENTS

RESULT 1	Q9UK79	PRELIMINARY;	PRT;	419 AA.
ID	Q9UK79			
AC	Q9UK79;			
DT	01-MAY-2000	(TREMBLrel. 13, Created)		
DT	01-MAR-2001	(TREMBLrel. 16, Last sequence update)		
DT	01-JUN-2002	(TREMBLrel. 21, Last annotation update)		
DE	Hersatin.			
GN	HER-2.			
OS	Homo sapiens			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Elasmobranchii;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo			
OX	"The HER-2/neu receptor tyrosine kinase gene encodes a secreted protein."			
RN	[1]			
RR	SEQUENCE FROM N.A.			
RX	MEDLINE=9941551; PUBMED=10485918;			
RA	Doherty J.K., Bond C., Jardim A., Adelman J.P., Clinton G.M.;			
RT	NCBL_TaxID=9606;			
RT	"The HER-2/neu receptor tyrosine kinase gene encodes a secreted protein."			
RT	Natl. Acad. Sci. U.S.A. 96:10869-10874(1999).			
RN	[2]			
RR	SEQUENCE FROM N.A.			
RA	Doherty J.K., Clinton G.M., Adelman J.P., Evans A.J., Henner W.D.;			
RL	Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AF17761; AA55609; 2;			
DR	InterPro: IPR000494; EGFR-L domain.			
DR	InterPro: IPR00174; Furin-like.			
DR	Furin; PF00757; Furin-like; 1.			
PFam;	PF01030; Recep,L domain; 1.			
DR	SMART; SM00261; FU; 1.			
SQ	SEQUENCE 419 AA; 45472 MW; FEC1BE347E2D30C CRC64;			

Query Match 19.8%; Score 83; DB 4; Length 419;  
Best Local Similarity 10.0%; Pre 5.6e-77;  
Matches 83; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Q9UK79 CKGSCRGCGSSEDCOSLTRIVCAGGCARKGKPLPDDCCHEQCAANGTGKHSCLACLHF 258  
Q9WVF5 cricetulus

Db 199 CKGSCRGCGSSEDCOSLTRIVCAGGCARKGKPLPDDCCHEQCAANGTGKHSCLACLHF 258  
Q9UVJ2 oreochromis

OY	259	NHSICELHC PALVNTYNTDFES	281
Db	259	NHSICELHC PALVNTYNTDFES	281
RESULT 2			
ID	Q9BG66	PRELIMINARY;	PRT; 149 AA.
AC	Q9BG66;		
DT	01-JUN-2001 (TREMBREL 17; Last sequence update)		
DR	01-OCT-2001 (TREMBREL 18; Last annotation update)		
DE	Receptor tyrosine kinase Erbb2 (Fragment).		
DR	Oryctolagus cuniculus (Rabbit).		
OS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.		
OC			
OX	NCBI_TAXID=9986;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RA	Klonisch T., Wolf P., Hombach-Klonisch S., Vogt S., Ruechenhoff A., Tetens F., Fischer B.;		
RT	"Erbb genes and epidermal growth factor- (EGF-) like ligands in the peri-implantation rabbit uterus and blastocyst"; Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.		
RL	EMBL; AE333171; AAK14371.1; -.		
DR	InterPro; IPR002174; Furin-like.		
DR	Pfam; PF0057; Furin-like; 1.		
DR	SMART; SM00261; FU; 2.		
KW	Kinase.		
FT	NON_TER 1 1		
FT	NON_TER 149 149 AA; 16240 MW; 7CB3792A54FC49BA CRC64;		
SQ	SEQUENCE FROM N.A.		
Query Match	12.2%; Score 51; DB 6; Length 149;		
Best Local Similarity	100.0%;	Pred. No. 2.7e-44;	
Matches	51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
DR	01-JAN-1998 (TREMBREL 05; Last sequence update)		
QY	231 LPTDCCEQCAAGCTGPKHSCLACLHFHNHSICELHC PALVNTYNTDFES	281	
Db	74 LPTDCCEQCAAGCTGPKHSCLACLHFHNHSICELHC PALVNTYNTDFES	124	
RESULT 3			
ID	Q18735	PRELIMINARY;	PRT; 1259 AA.
AC	Q18735;		
DT	01-JAN-1998 (TREMBREL 05; Last sequence update)		
DR	01-JAN-1998 (TREMBREL 05; Last sequence update)		
DE	01-JUN 2002 (TREMBREL 21; Last annotation update)		
OS	Canis familiaris (Dog).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.		
OX	NCBI_TAXID=9615;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RA	Yokota H.;		
RT	"cDNA cloning of erb-2 from canine mammary gland.";		
DR	Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; AB008451; BAA23127.1; -.		
DR	HSRP; PI1362; IFGK.		
DR	InterPro; IPR002048; EP-hand.		
DR	InterPro; IPR00494; EGFR_L_domain.		
DR	InterPro; IPR000119; Euk-pkinase.		
DR	InterPro; IPR002174; Furin-like.		
DR	InterPro; IPR00245; Tyr_pkinase.		
DR	InterPro; IPR004019; YLP motif.		
DR	Pfam; PF0057; Furin-like; 1.		
DR	Pfam; PF00659; pkinase; 1.		
DR	Pfam; PF01030; Recop_L_domain; 2.		
DR	Pfam; PF02257; YLP; 2.		
DR	ProDom; PD000001; Euk_pkinase; 1.		
RESULT 4			
ID	Q9BSE0	PRELIMINARY;	PRT; 478 AA.
AC	Q9BSE0;		
DT	01-MAR-2001 (TREMBREL 16; Created)		
DR	01-OCT-2001 (TREMBREL 18; Last sequence update)		
DE	Epidermal growth factor receptor related protein.		
GN	EREP.		
OS	Rattus norvegicus (Rat).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.		
OX	NCBI_TAXID=10116;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE-GASTRO-DUODENAL MUCOUS;		
RA	Yu Y., Moshtagh J.A., Majumdar A.P.N.;		
RT	"Cloning of a novel EGF receptor regulator of EGFR"; Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.		
RL	EMBL; AF187818; AAGI1037.2; -.		
DR	InterPro; IPR00049; EGFR_L_domain.		
DR	InterPro; IPR002174; Furin-like.		
DR	Pfam; PF00757; Furin-like; 1.		
DR	Pfam; PF01030; Recop_L_domain; 2.		
KW	Receptor.		
SQ	SEQUENCE FROM N.A.		
Query Match	2.1%; Score 9; DB 11; Length 478;		
Best Local Similarity	100.0%;	Pred. No. 2;	
Matches	9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
DR	01-JUN 2002 (TREMBREL 21; Last annotation update)		
QY	239 QCAAGCNGP 247		
Db	235 QCAGCNGP 243		
RESULT 5			
ID	Q9BSE6	PRELIMINARY;	PRT; 527 AA.
AC	Q9BSE6;		
DT	01-NOV-1996 (TREMBREL 01; Created)		
DR	01-NOV-1996 (TREMBREL 01; Last sequence update)		
DT	01-JUN 2002 (TREMBREL 21; Last annotation update)		
DE	EGF/TGF-alpha receptor precursor.		
GN	C-ERBB.		
OS	Gallus gallus (Chicken).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.		
OX	NCBI_TAXID=9031;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
DR	MEDLINE-92123214; PubMed=1732751;		
RX	Flickinger T.W., Maihile N.J., Kang H.-J.;		

RT	"An alternatively processed mRNA from the avian c-erbB gene encodes a soluble, truncated form of the receptor that can block ligand-dependent transformation.";
RT	Mol. Cell. Biol. 12:883-893(1992).
RL	EMBL; M7637; AAA48759.1; -.
DR	InterPro; IPR00494; EGFR_L-domain.
DR	Pfam; PF00757; Furin-like; 1.
DR	Pfam; PF01030; Recep_L-domain; 2.
DR	SMART; SMM0261; FU; 2.
KW	Receptor; Signal.
FT	POTENTIAL_SIGNAL.
FT	RECEPTOR; Signal.
SEQUENCE	527 AA; 58353 MW; 764564ABCC095298 CRC64;
CHAIN	29
CHAIN	1 28
CHAIN	29 527 EGF/TGF-ALPHA RECEPTOR.
CHAIN	1 28 POTENTIAL_SIGNAL.
RQSH2	OPPSH2 PRELIMINARY; PRT; 599 AA.
ID	OPPSH2;
AC	OPPSH2;
DT	01-MAY-2000 (TREMBLrel. 13, Created)
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT	01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE	Epidermal growth factor receptor (fragment).
OS	Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Gallus gallus (Chicken).
OC	Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus; Gallus.
OC	NCBI_TAXID=9031;
RN	[1]
RP	SEQUENCE FROM N_A.
RX	MEDLINE=94020816; PubMed=8414496;
RA	Callaghan T., Antczak M., Flickinger T., Raines M., Myers M., Kung H.-J.
RA	"A complete description of the EGFR-receptor exon structure: Implication in oncogenic activation and domain evolution.";
RA	Oncoogene 8:2939-2948(1993).
DR	InterPro; IPR00494; EGFR_L-domain.
DR	InterPro; IPR02174; Furin-like.
DR	Pfam; PF00757; Furin-like; 2.
DR	Pfam; PF01030; Recep_L-domain; 2.
DR	SMART; SM00261; FU; 3.
SQ	SEQUENCE 599 AA; 66363 MW; FEAB46D293D991BD CRC64;
RESULT	7
Query Match	2.1%; Score 9; DB 13; Length 599;
Best Local Similarity	100.0%; Pred. No. 2,4;
Matches	9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	239 QCAAGCTGP 247
Db	393 QCAAGCTGP 401
Q9PSH2	OPPSH2 PRELIMINARY; PRT; 599 AA.
ID	OPPSH2;
AC	OPPSH2;
DT	01-NOV-1999 (TREMBLrel. 12, Created)
DT	01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DE	Epidermal growth factor receptor (Epidermal growth factor receptor isoform 3).
DE	EGFR.
CN	Mus musculus (Mouse).
OS	Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC	NCBI_TAXID=10090;
RN	[1]
RP	SEQUENCE FROM N_A.
RC	STRAIN=C57BL/6J; TISSUE=LIVER;
RA	Reiter J.L., Threadgill D.W., Danielson A.J., Schell C., Lampland A.L., Balsubramanian S., Crossley T.O., Magnuson T.R., Maahle N.J.
RA	"Alternative Transcripts from the Human and Mouse EGFR Genes Encode Carboxy-terminal Truncated Receptors.";
RA	Submitted (JAN-1999) to the EMBL/GenBank/DDBJ databases.
RN	[2]
RP	SEQUENCE FROM N_A.
RC	STRAIN=C3H/101, 129/SVJ, AND 129/SVETAC;
RA	Reiter J.L., Threadgill D.W., Eley G.D., Strunk K.E., Danielson A.J., Sinclair C.S., Pearseill R.S., Green P.J., Yee D., Lampland A.L., Balsubramanian S., Crossley T.O., Magnuson T.R., James C.D., Maahle N.J.
RA	"Comparative genomic sequence analysis and isolation of human and mouse alternative EGFR transcripts encoding truncated receptor isoforms.";
RA	Submitted (JUN-2000) to the EMBL/GenBank/DDBJ databases.
RN	[3]
RP	SEQUENCE FROM N_A.
RC	STRAIN=C57BL/6J; TISSUE=LIVER;
RC	MEDLINE=21085660; PubMed=11218851;
RT	NCBL_TaxID=10090;
RN	[1]
RP	SEQUENCE FROM N_A.
RC	STRAIN=C3H/101, 129/SVJ, AND 129/SVETAC;
RA	Reiter J.L., Threadgill D.W., Strunk K.E., Danielson A.J., Sinclair C.S., Pearseill R.S., Green P.J., Yee D., Lampland A.L., Balsubramanian S., Crossley T.O., Magnuson T.R., James C.D., Maahle N.J.;
RA	"Comparative genomic sequence analysis and isolation of human and mouse alternative EGFR transcripts encoding truncated receptor isoforms.";
RA	Submitted (JUN-2000) to the EMBL/GenBank/DDBJ databases.
RT	"Comparative genomic sequence analysis and isolation of human and mouse alternative Egfr transcripts encoding truncated receptor isoforms.";
RT	Submitted (JUN-2000) to the EMBL/GenBank/DDBJ databases.
RT	MAHLE N.J.;
DR	EMBL; AF215364; AAG2046.1; JOINED.
DR	MGD; MGI:95294; Egfr.
DR	InterPro; IPR00345; CYTC_heme_bind.
DR	InterPro; IPR00494; EGFR_L-domain.
DR	InterPro; IPR0074; Furin-like.
DR	Pfam; PF00757; Furin-like; 1.
DR	Pfam; PF01030; Recep_L-domain; 2.
DR	SMART; SMM0261; FU; 4.
DR	PROSITE; PS00190; CYTOCHROME_C; UNKNOWN_1.
KW	Receptor.
SQ	SEQUENCE 643 AA; DEF22002C8491B1 CRC64;
RESULT	8
Q9WVF5	Q9WVF5 PRELIMINARY; PRT; 655 AA.
ID	Q9WVF5;
AC	Q9WVF5;
DT	01-NOV-1999 (TREMBLrel. 12, Created)
DT	01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE	Epidermal growth factor receptor (Epidermal growth factor receptor isoform 3).
DE	EGFR.
CN	Mus musculus (Mouse).
OS	Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC	NCBI_TAXID=10090;
RN	[1]
RP	SEQUENCE FROM N_A.
RC	STRAIN=C57BL/6J; TISSUE=LIVER;
RA	Reiter J.L., Threadgill D.W., Danielson A.J., Schell C., Lampland A.L., Balsubramanian S., Crossley T.O., Magnuson T.R., Maahle N.J.
RA	"Alternative Transcripts from the Human and Mouse EGFR Genes Encode Carboxy-terminal Truncated Receptors.";
RA	Submitted (JAN-1999) to the EMBL/GenBank/DDBJ databases.
RN	[2]
RP	SEQUENCE FROM N_A.
RC	STRAIN=C3H/101, 129/SVJ, AND 129/SVETAC;
RA	Reiter J.L., Threadgill D.W., Eley G.D., Strunk K.E., Danielson A.J., Sinclair C.S., Pearseill R.S., Green P.J., Yee D., Lampland A.L., Balsubramanian S., Crossley T.O., Magnuson T.R., James C.D., Maahle N.J.
RA	"Comparative genomic sequence analysis and isolation of human and mouse alternative EGFR transcripts encoding truncated receptor isoforms.";
RA	Submitted (JUN-2000) to the EMBL/GenBank/DDBJ databases.
RN	[3]
RP	SEQUENCE FROM N_A.
RC	STRAIN=C57BL/6J; TISSUE=LIVER;
RC	MEDLINE=21085660; PubMed=11218851;

RA kawai J., Shinagawa A., Shibusawa K., Yoshino M., Itoh M., Ishii Y.,  
 RA Akirawa T., Hira J., Fukunishi Y., Kono H., Radchi J., Fukuda S.,  
 RA Aizawa K., Izawa M., Nishi K., Kiyoawara H., Kondo S., Yamakawa T.,  
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
 RA Fleischmann R., Gaasterland T., Gissi C., King B., Kochiwa H.,  
 RA Schirini P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,  
 RA Sakai L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
 RA Sakai L., Ohido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamura M., Lee N.H.,  
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombberts P.,  
 RA Nordone P., Ringwald M., Rodriguez I., Sakamoto N.,  
 RA Sasaki H., Sato K., Schoenbach C., Seiya T., Storch K.-F.,  
 RA Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Wittaker C., Wimling L.,  
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,  
 RA Hayashizaki Y.,  
 RT "Functional annotation of a full-length mouse cDNA collection.";  
 RL Nature 409:685-690(2001).  
 DR EMBL; AF124513; AAD41149.1; -.  
 DR EMBL; AF275366; AAG28047.1; JOINED.  
 DR EMBL; AF275364; AAG28047.1; JOINED.  
 DR EMBL; AF275365; AAG28047.1; JOINED.  
 DR EMBL; AF275366; AAG28047.1; JOINED.  
 DR EMBL; AK004944; BAB23688.1; -.  
 DR EMBL; AK004933; BAB23641.1; -.  
 DR EMBL; AF275366; AAG28047.1; -.  
 DR InterPro; IPR000494; EGFR\_L\_domain.  
 DR Pfam; PF00757; Furin-like.  
 DR Pfam; PF00757; Furin-like.  
 DR Pfam; PF01030; Recep\_Ldomain; 2.  
 DR SMART; SM00261; FU: 3.  
 DR SMART; SM00219; TYRK; 1.  
 DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
 DR ATP-binding; Receptor; Transf erase; Tyrosine-Protein kinase.  
 DR Sequence 1209 AA; 134891 MW; 96FEETF6CCIB773 CRC64;  
 DR Kw  
 SQ Query Match 2.1%; Score 9; DB 11; Length 1209;  
 Best Local Similarity 100.0%; Pred. No. 4.5;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 DR QY 239 QCAAGCTGP 247  
 DR Db 235 QCAAGCTGP 243  
 RESULT 9  
 09QX70 PRELIMINARY; PRT; 1209 AA.  
 ID 09QX70 PRELIMINARY; PRT; 1209 AA.  
 AC 09QX70:  
 DT 01-MAY-2000 (TREMBL; 13, Created)  
 DT 01-MAY-2000 (TREMBL; 13, Last sequence update)  
 DT 01-JUN-2002 (TREMBL; 21, Last annotation update)  
 DE Epidermal growth factor receptor isoform 1.  
 DE EGFR.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CCH/101, 129/SVJ, AND 129/SVETVAC;  
 RA Reiter J.L., Threadgill D.W., Eley G.D., Strunk K.E., Danielsen A.J.,  
 RA Sinclair C.S., Pearsall R.S., Green P.J., Yee D., Lampland A.L.,  
 RA Balasubramanian S., Crossley T.O., Magnuson T.R., James C.D.,  
 RA Maihle N.J.;  
 RA "Comparative genomic sequence analysis and isolation of human and  
 RT mouse alternative Egfr transcripts encoding truncated receptor  
 RT isoforms";  
 RT Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.  
 RL Sequence FROM N.A.  
 RN [2]  
 RP STRAIN=C57BL/6J;  
 RC STRAIN=C57BL/6J;  
 RA Reiter J.L., Threadgill D.W., Eley G.D., Strunk K.E., Danielsen A.J.,  
 RA Balasubramanian S., Crossley T.O., Magnuson T.R., James C.D.,  
 RA Schehl C., Pearsall R.S., Green P.J., Yee D., Lampland A.L.,  
 RA Maihle N.J.;  
 RA "Comparative genomic sequence analysis and isolation of human and  
 RT mouse alternative Egfr transcripts encoding truncated receptor  
 RT isoforms";  
 RT Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.  
 RL EMBL; AF275366; AAG28045.1; -.  
 DR EMBL; AF275364; AAC28045.1; JOINED.  
 DR EMBL; AF275365; AAG28045.1; JOINED.  
 DR EMBL; AF275367; AAG28045.1; -.  
 DR EMBL; P11362; IFGK-  
 DR MGD; MGI:95294; Egfr.  
 DR

DR InterPro; IPR000345; CYC\_heme\_bind.  
 DR InterPro; IPR000494; EGFR\_L\_domain.  
 DR InterPro; IPR000719; Euk\_pk kinase.  
 DR InterPro; IPR002174; Furin-like.  
 DR InterPro; IPR002290; Ser\_thr\_pk kinase.  
 DR InterPro; IPR001245; Tyr\_pk kinase.  
 DR Pfam; PF000757; Furin-like; 1.  
 DR Pfam; PF00069; Pkinase; 1.  
 DR Pfam; PF01030; Recp\_L\_domain; 2.  
 DR PRINTS; PR00109; TYRKINASE.  
 DR PRODOM; PD00001; Euk\_pk kinase; 1.  
 DR SMART; SM00261; FU; 5.  
 DR SMART; SM00220; S\_TK\_C; 1.  
 DR PROSITE; PS00190; CTOCHROME\_C; UNKNOWN\_1.  
 DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE; PS0109; PROTEIN\_KINASE\_TIR; 1.  
 KW ATP-binding; Receptor; Transferase.  
 SQ SEQUENCE 1210 AA; 134840 MW; 62CDD021C9DE32E18 CRC64;

Query Match 2.1%; Score 9; DB 11; Length 1210;  
 Best Local Similarity 100.0%; Pred No. 4.5;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 239 QCAAGCTGP 247  
 Db 235 QCAAGCTGP 243

RESULT 11

063365 PRELIMINARY; PRT; 48 AA.  
 ID 063365; AC 063365; DT 01-NOV-1996 (TREMBLrel. 01, created)  
 DT 01-NOV-1996 (TREMBLrel. 01, last sequence update)  
 DT 01-MAR-2002 (TREMBLrel. 20, last annotation update)  
 DE Transforming oncogene NEU (Fragment).  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 NCBL-TAXID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=91061737; PUBMED=2122292;  
 RA Suen T.C.; Hung M.C.;  
 RT "Multiple cis- and trans-acting elements involved in regulation of the neu gene";  
 RL Mol. Cell. Biol. 10:6306-6315(1990).  
 RT [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=91172172; PUBMED=1672439;  
 RA Yan D.H.; Hung M.C.;  
 RT "Identification and characterization of a novel enhancer for the rat rt neu promoter";  
 RL Mol. Cell. Biol. 11:1875-1882(1991).  
 DR EMBL; M61004; AAA41686.1;  
 KW Oncogene; Transforming protein.  
 FT NON\_TER 48  
 SQ SEQUENCE 48 AA; 4966 MW; 72A054B1E13F1129 CRC64;

Query Match 1.9%; Score 8; DB 11; Length 48;  
 Best Local Similarity 100.0%; Pred No. 2.9;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 12

09QZ16 PRELIMINARY; PRT; 63 AA.  
 Qy 12 LLLLPGG 19  
 Db 15 LLLLPGG 22

RESULT 13

09QZ15 PRELIMINARY; PRT; 63 AA.  
 ID 09QZ15; AC 09QZ15; DT 01-MAY-2000 (TREMBLrel. 13, created)  
 DT 01-MAY-2000 (TREMBLrel. 13, last sequence update)  
 DT 01-MAY-2000 (TREMBLrel. 13, last sequence update)  
 DT 01-MAY-2000 (TREMBLrel. 13, last annotation update)  
 DE Carbonyl reductase Isoform I (Fragment).  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 NCBL-TAXID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX STRAIN=WISTAR; TISSUE=OVARY;  
 RA Esprey L.L.; Yoshioka S.; Russell D.; Ujioka T.; Vladu B.; Skelsey M.;  
 RA Fujii S.; Okamura H.; Richards J.S.;  
 RT "Characterization of Ovarian Carbonyl Reductase Gene Expression during Ovulation in the Gonadotropin-Primed Immature Rat";  
 RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AFI81956; AAF03395.1;  
 FT NON\_TER 1  
 SQ SEQUENCE 63 AA; 6916 MW; D5964BC1E299A2F4 CRC64;

Query Match 1.9%; Score 8; DB 11; Length 63;  
 Best Local Similarity 100.0%; Pred No. 3.6;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 14

008558 PRELIMINARY; PRT; 277 AA.  
 ID 008558; AC 008558; DT 01-JUL-1997 (TREMBLrel. 04, created)  
 DT 01-JUL-1997 (TREMBLrel. 04, last sequence update)  
 DT 01-DIC-2001 (TREMBLrel. 19, last annotation update)  
 DE Inducible carbonyl reductase.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OC NCBI\_TaxID=10116;  
 RN [1]  
 RP  
 RQ SEQUENCE FROM N.A.  
 RC STRAIN=WISTER; TISSUE=OVARY;  
 RX MEDLINE=9716735; PubMed=9115353;  
 RA Aoki H.; Okada T.; Mizutani T.; Numata Y.; Minegishi T.; Miyamoto K.;  
 RT "Identification of two closely related genes, inducible and  
 noninducible carbonyl reductases in the rat ovary.";  
 RL Biochem. Biophys. Res. Commun. 230:518-523 (1997).  
 CC -1 SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES  
 (SPR) FAMILY.  
 DR EMBL; D89069; BAA9007\_1; -.  
 DR IISPP; P50162; IAEI.  
 DR InterPro; IPR002198; ADH\_short.  
 DR PRAM; PR00106; adh\_short; 1.  
 DR PROSITE; PS00061; ADH\_SHORT; UNKNOWN\_1.  
 KW Oxidoreductase.  
 SQ SEQUENCE 277 AA; 30654 MW; DDA015D1B71673A7 CRC64;  
 Query Match 1.9%; Score 8; DB 11; Length 277;  
 Best Local Similarity 100.0%; Pred. No. 13;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 13 LALUPPGA 20  
 |||||  
 Db 254 LALUPPGA 261

## RESULT 15

Q924V3 PRELIMINARY; PRT; 277 AA.  
 ID Q924V3:  
 AC 0924V3:  
 DT 01-DEC-2001 (TREMBREL, 19, Created)  
 DT 01-MAR-2002 (TREMBREL, 20, Last sequence update)  
 DE Carbonyl reductase 1.  
 GN CTCR1.  
 OS Cricetulus griseus (Chinese hamster).  
 OC Mammalia; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Eukaryota; Metazoa; Rodentia; Sciurognathi; Muridae; Cricetinae;  
 OC Cricetulus.  
 OX NCBI\_TaxID=10029;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Terada T.; Sugimura Y.; Nakamura K.;  
 RT "Chinese hamster carbonyl reductase.";  
 RL Submitted (MAY-2000) to the EMBL/GenBank/DDBJ databases.  
 CC -1 SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES  
 (SPR) FAMILY.  
 DR EMBL; AB04541; BAB62840\_1; -.  
 DR InterPro; IPR002198; ADH\_short.  
 DR PRAM; PR00106; adh\_short; 1.  
 DR PROSITE; PS00061; ADH\_SHORT; UNKNOWN\_1.  
 KW Oxidoreductase.  
 SQ SEQUENCE 277 AA; 30498 MW; 2F7EB76E2B109A31 CRC64;

Query Match 1.9%; Score 8; DB 11; Length 277;  
 Best Local Similarity 100.0%; Pred. No. 13;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 13 LALUPPGA 20  
 |||||  
 Db 254 LALUPPGA 261

Search completed: January 14, 2003, 17:14:25  
 Job time : 49.2751 secs

Run on: January 14, 2003, 17:14:40 ; Search time 14.3032 seconds  
 (without alignments)  
 568.337 Million cell updates/sec

Title: US-09-506-079E-2  
 Perfect score: 419  
 Sequence: 1 MELAALCRWGLLLALLPPGA.....VGRGXDPDAHVAVLSRYEG 419

Scoring table: OLIGO Gapop 60.0 , Gapext 60.0

Searched: 118974 seqs; 19410157 residues

Word size : 0

Total number of hits satisfying chosen parameters: 118974

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : Published\_Applications\_MA:\*

1: /cgn2\_6/ptodata/2/pubpea/US08\_NEW\_PUB\_pep:\*

2: /cgn2\_6/ptodata/2/pubpea/PCT\_NEW\_PUB\_pep:\*

3: /cgn2\_6/ptodata/2/pubpea/US06\_NEW\_PUB\_pep:\*

4: /cgn2\_6/ptodata/2/pubpea/US07\_NEW\_PUB\_pep:\*

5: /cgn2\_6/ptodata/2/pubpea/US07\_PUBCOMB\_pep:\*

6: /cgn2\_6/ptodata/2/pubpea/US07\_PUBCOMB\_pep:\*

7: /cgn2\_6/ptodata/2/pubpea/PCRS\_PUBCOMB\_pep:\*

8: /cgn2\_6/ptodata/2/pubpea/US08\_PUBCOMB\_pep:\*

9: /cgn2\_6/ptodata/2/pubpea/US09\_NEW\_PUB\_pep:\*

10: /cgn2\_6/ptodata/2/pubpea/US10\_PUBCOMB\_pep:\*

11: /cgn2\_6/ptodata/2/pubpea/US10\_NEW\_PUB\_pep:\*

12: /cgn2\_6/ptodata/2/pubpea/US10\_PUBCOMB\_pep:\*

13: /cgn2\_6/ptodata/2/pubpea/US10\_NEW\_PUB\_pep:\*

14: /cgn2\_6/ptodata/2/pubpea/US10\_PUBCOMB\_pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

**SUMMARIES**

Result No.	Score	Query Match length	DB ID	Description
1	83	19.8	289	10 US-09-821-883-23
2	83	19.8	479	10 US-09-821-883-5
3	83	19.8	555	10 US-09-821-883-1
4	83	19.8	564	10 US-09-821-883-3
5	83	19.8	645	10 US-09-821-881-1
6	83	19.8	653	9 US-09-854-356-3
7	83	19.8	690	10 US-09-821-883-2
8	83	19.8	697	10 US-09-821-883-4
9	83	19.8	712	9 US-09-854-356-7
10	83	19.8	919	9 US-09-854-356-6
11	83	19.8	1255	9 US-09-854-356-2
12	83	19.8	1255	9 US-09-854-356-1
13	83	19.8	1255	9 US-09-854-356-2
14	83	19.8	1255	10 US-09-811-115-3
15	83	19.8	1255	10 US-09-811-115-3
16	51	12.2	654	9 US-09-854-356-8
17	51	12.2	1256	9 US-09-854-356-2
18	51	12.2	1256	9 US-09-870-759-18
19	9.8	0	0	US-09-854-356-14

**ALIGNMENTS**

RESULT 1  
 US-09-821-883-23  
 ; Sequence 23, Application US/09821883  
 ; Patent No. US20020061310A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Laus, Reiner  
 ; APPLICANT: Vidovic, Damir  
 ; APPLICANT: Graddis, Thomas  
 ; TITLE OF INVENTION: Compositions and Methods for Dendritic FILE REFERENCE: 7636-0022-30  
 ; CURRENT APPLICATION NUMBER: US/09/821, 883  
 ; CURRENT FILING DATE: 2001-03-30  
 ; PRIOR APPLICATION NUMBER: US 60/193, 504  
 ; NUMBER OF SEQ ID NOS: 30  
 ; SOFTWARE: FASTSEQ for Windows Version 4.0  
 ; SEQ ID NO: 23  
 ; LENGTH: 289  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; US-09-821-883-23

Query Match Score: 83; DB ID: 10; Length: 289;  
 Best Local Similarity: 19.8%; Pred. No.: 2.7e-68;  
 Matches: 83; Conservative: 0; Mismatches: 0; Indels: 0; Gaps: 0;

QY 199 CKGSRCWGESSEDQLTRIVCAGGCARKKGPLPDCCHFCQACGTGKHSDCLACLR 258  
 Db 178 CKGSRCWGESSEDQLTRIVCAGGCARKKGPLPDCCHFCQACGTGKHSDCLACLR 237

QY 259 NHSGICELHCPALVYNTDFFES 281  
 Db 238 NHSGICELHCPALVYNTDFFES 260

RESULT 2  
 US-09-821-883-5  
 ; Sequence 5, Application US/09821883  
 ; Sequence 4, Appli  
 Sequence 7, Appli  
 Sequence 6, Appli  
 Sequence 2, Appli  
 Sequence 1, Appli  
 Sequence 2, Appli  
 Sequence 9, Appli  
 Sequence 3, Appli  
 Sequence 8, Appli  
 Sequence 2, Appli  
 Sequence 118, Appli  
 Sequence 14, Appli  
 ; APPLICANT: Laus, Reiner

RESULT 4  
US-09-821-883-3  
; Sequence 3, Application US/09821883

Query Match 19.8%; Score 83; DB 10; Length 555;  
Best Local Similarity 100.0%; Pred. No. 4.Be-68;  
Matches 83; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 199 CKGSRMGESSEDCQSITRTVAGGGCARCKGKGLPTDCHEQCAAGCTGPKHSCLACLHF 258  
Db 218 CKGSRMGESSEDCQSITRTVAGGGCARCKGKGLPTDCHEQCAAGCTGPKHSCLACLHF 277

Qy 259 NHSGICELHCPALVTNTDFFES 281  
Db 278 NHSGICELHCPALVTNTDFFES 300

RESULT 3  
US-09-821-883-1  
; Sequence 1, Application US/09821883  
; Patent No. US2002061310A1

GENERAL INFORMATION:  
; APPLICANT: Laus, Reiner  
; APPLICANT: Vidovic, Danir  
; APPLICANT: Graddis, Thomas  
; TITLE OF INVENTION: Compositions and Methods for Dendritic  
; TITLE OF INVENTION: Cell-Based Immunotherapy

CURRENT APPLICATION NUMBER: US/09/821-883  
PRIORITY FILING DATE: 2001-03-30  
NUMBER OF SEQ ID NOS: 30  
SOFTWARE: FastSEQ for Windows Version 4.0  
SEQ ID NO 5  
LENGTH: 479  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
; OTHER INFORMATION: HER300\*-rGM-CSF construct

Query Match 19.8%; Score 83; DB 10; Length 479;  
Best Local Similarity 100.0%; Pred. No. 4.2e-68;  
Matches 83; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 199 CKGSRMGESSEDCQSITRTVAGGGCARCKGKGLPTDCHEQCAAGCTGPKHSCLACLHF 258  
Db 218 CKGSRMGESSEDCQSITRTVAGGGCARCKGKGLPTDCHEQCAAGCTGPKHSCLACLHF 277

Qy 259 NHSGICELHCPALVTNTDFFES 281  
Db 278 NHSGICELHCPALVTNTDFFES 300

RESULT 4  
US-09-821-883-1

Query Match 19.8%; Score 83; DB 10; Length 564;  
Best Local Similarity 100.0%; Pred. No. 4.8e-68;  
Matches 83; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 199 CKGSRMGESSEDCQSITRTVAGGGCARCKGKGLPTDCHEQCAAGCTGPKHSCLACLHF 258  
Db 218 CKGSRMGESSEDCQSITRTVAGGGCARCKGKGLPTDCHEQCAAGCTGPKHSCLACLHF 277

Qy 259 NHSGICELHCPALVTNTDFFES 281  
Db 278 NHSGICELHCPALVTNTDFFES 300

RESULT 5  
US-09-921-161-1  
; sequence 1, Application US/09921161  
; Patent No. US2002080662A1

GENERAL INFORMATION:  
; APPLICANT: Ralph, Peter  
; APPLICANT: Peter  
; TITLE OF INVENTION: ANALYTICAL METHOD  
FILE REFERENCE: GEMENT\_066A

CURRENT APPLICATION NUMBER: US/09/921-161  
CURRENT FILING DATE: 2001-08-01  
PRIORITY FILING DATE: 2000-08-15  
NUMBER OF SEQ ID NOS: 1  
SOFTWARE: FastSEQ for Windows Version 4.0  
SEQ ID NO 1  
LENGTH: 645  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
; OTHER INFORMATION: HER500 construct

Query Match 19.8%; Score 83; DB 10; Length 645;  
Best Local Similarity 100.0%; Pred. No. 5.4e-68;  
Matches 83; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 199 CKGSRMGESSEDCQSITRTVAGGGCARCKGKGLPTDCHEQCAAGCTGPKHSCLACLHF 258  
Db 218 CKGSRMGESSEDCQSITRTVAGGGCARCKGKGLPTDCHEQCAAGCTGPKHSCLACLHF 277

Qy 259 NHSGICELHCPALVTNTDFFES 281  
Db 278 NHSGICELHCPALVTNTDFFES 300

RESULT 6  
US-09-854-356-3  
; Sequence 3, Application US/09854356  
; Patent No. US2002017756A1

GENERAL INFORMATION:

APPLICANT: Cheever, Martin A.  
 APPLICANT: Gheysen, Dirk  
 APPLICANT: Corixa Corporation  
 APPLICANT: SmithKline Beecham Biologicals S. A.  
 TITLE OF INVENTION: HER-2/neu Fusion Proteins  
 FILE REFERENCE: 014058-009810PC

CURRENT APPLICATION NUMBER: US 09/854,356  
 CURRENT FILING DATE: 2001-05-09  
 PRIOR APPLICATION NUMBER: US 09/493,480  
 PRIOR FILING DATE: 1999-01-29  
 NUMBER OF SEQ ID NOS: 26  
 SOFTWARE: PatentIn Ver. 2.1  
 SEQ ID NO: 3  
 LENGTH: 653  
 TYPE: PRT  
 ORGANISM: Homo sapiens

FEATURE:  
 OTHER INFORMATION: extracellular domain (ECD) of human HER-2/neu

Query Match, Best Local Similarity 19.8%; Score 83; DB 9; Length 653;  
 Matches 83; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 199 CKGSRCWGSSSEDCQSLTRIVCAGGCARCKGPLPTDCHEQCAAGCTGPKHSCLACIHF 258  
 Db 199 CKGSRCWGSSSEDCQSLTRIVCAGGCARCKGPLPTDCHEQCAAGCTGPKHSCLACIHF 258

QY 259 NHSGICELICPALVYNTDFFES 281  
 Db 259 NHSGICELICPALVYNTDFFES 281

RESULT 7  
 US-09-821-883-2  
 ; Sequence 2, Application US/09821883  
 ; Patent No. US20020061310A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Graddis, Thomas  
 ; APPLICANT: Vidovic, Damir  
 ; APPLICANT: Laus, Reiner  
 ; APPLICANT: Graddis, Thomas  
 ; TITLE OF INVENTION: Compositions and Methods for Dendritic  
 ; TITLE OF INVENTION: Cell-Based Immunotherapy  
 ; FILE REFERENCE: 7636-0022.30  
 ; CURRENT APPLICATION NUMBER: US 09/821,883  
 ; CURRENT FILING DATE: 2001-03-30  
 ; PRIOR APPLICATION NUMBER: US 60/193,504  
 ; PRIOR FILING DATE: 2000-03-30  
 ; NUMBER OF SEQ ID NOS: 30  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO: 4  
 ; LENGTH: 697  
 ; TYPE: PRT  
 ; ORGANISM: Artificial Sequence

FEATURE:  
 OTHER INFORMATION: HER500\*-rGM-CSF construct

Query Match, Best Local Similarity 19.8%; Score 83; DB 10; Length 697;  
 Matches 83; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 199 CKGSRCWGSSSEDCQSLTRIVCAGGCARCKGPLPTDCHEQCAAGCTGPKHSCLACIHF 258  
 Db 218 CKGSRCWGSSSEDCQSLTRIVCAGGCARCKGPLPTDCHEQCAAGCTGPKHSCLACIHF 277

QY 259 NHSGICELICPALVYNTDFFES 281  
 Db 278 NHSGICELICPALVYNTDFFES 300

RESULT 8  
 US-09-821-883-4  
 ; Sequence 4, Application US/09821883  
 ; Patent No. US20020061310A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Laus, Reiner  
 ; APPLICANT: Vidovic, Damir  
 ; APPLICANT: Graddis, Thomas  
 ; TITLE OF INVENTION: Compositions and Methods for Dendritic  
 ; TITLE OF INVENTION: Cell-Based Immunotherapy  
 ; FILE REFERENCE: 7636-0022.30  
 ; CURRENT APPLICATION NUMBER: US 09/821,883  
 ; PRIOR APPLICATION NUMBER: US 60/193,504  
 ; NUMBER OF SEQ ID NOS: 30  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO: 4  
 ; LENGTH: 697  
 ; TYPE: PRT  
 ; ORGANISM: Artificial Sequence

FEATURE:  
 OTHER INFORMATION: HER500\*-rGM-CSF construct

Query Match, Best Local Similarity 19.8%; Score 83; DB 10; Length 697;  
 Matches 83; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 199 CKGSRCWGSSSEDCQSLTRIVCAGGCARCKGPLPTDCHEQCAAGCTGPKHSCLACIHF 258  
 Db 278 NHSGICELICPALVYNTDFFES 300

RESULT 9  
 US-09-854-356-7  
 ; Sequence 7, Application US/09854356  
 ; Patent No. US20020177567A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Cheever, Martin A.  
 ; APPLICANT: Gheysen, Dirk  
 ; APPLICANT: Corixa Corporation  
 ; APPLICANT: SmithKline Beecham Biologicals S. A.  
 ; TITLE OF INVENTION: HER-2/neu Fusion Proteins  
 ; FILE REFERENCE: 014058-009810PC  
 ; CURRENT APPLICATION NUMBER: US 09/854,356  
 ; CURRENT FILING DATE: 2001-05-09  
 ; PRIOR APPLICATION NUMBER: US 09/493,480  
 ; PRIOR FILING DATE: 2000-01-28  
 ; PRIOR APPLICATION NUMBER: US 60/117,976  
 ; PRIOR FILING DATE: 1999-01-29  
 ; NUMBER OF SEQ ID NOS: 26  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO: 7  
 ; LENGTH: 712  
 ; TYPE: PRT  
 ; ORGANISM: Artificial Sequence

FEATURE:  
 OTHER INFORMATION: Description of Artificial Sequence:fusion protein

Query Match, Best Local Similarity 19.8%; Score 83; DB 9; Length 712;  
 Matches 83; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 199 CKGSRCWGSSSEDCQSLTRIVCAGGCARCKGPLPTDCHEQCAAGCTGPKHSCLACIHF 258  
 Db 218 CKGSRCWGSSSEDCQSLTRIVCAGGCARCKGPLPTDCHEQCAAGCTGPKHSCLACIHF 277

QY 259 NHSGICELICPALVYNTDFFES 281  
 Db 278 NHSGICELICPALVYNTDFFES 300

QY 199 CKGSRCWGSSSEDCQSLTRIVCAGGCARCKGPLPTDCHEQCAAGCTGPKHSCLACIHF 258  
 Db 199 CKGSRCWGSSSEDCQSLTRIVCAGGCARCKGPLPTDCHEQCAAGCTGPKHSCLACIHF 258

QY 259 NHSGICELHCPALVNTDFES 281  
 ||||||| ||||| ||||| |||||  
 Sequence 6, Application US/09854356  
 Db 259 NHSGICELHCPALVNTDFES 281

RESULT 10  
 US-09-854-356-6  
 ; Sequence 6, Application US/09854356  
 ; Patent No. US20020177567A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Cheever, Martin A.  
 ; APPLICANT: Gheysen, Dirk  
 ; APPLICANT: Corixa Corporation  
 ; TITLE OF INVENTION: SmithKline Beecham Biologicals S. A.  
 ; FILE REFERENCE: 014058-009810PC  
 ; CURRENT APPLICATION NUMBER: US/09/854,356  
 ; CURRENT FILING DATE: 2001-05-09  
 ; PRIORITY APPLICATION NUMBER: US 09/493,480  
 ; PRIORITY FILING DATE: 1999-01-29  
 ; NUMBER OF SEQ ID NOS: 26  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 6  
 ; LENGTH: 919  
 ; TYPE: PRT  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: Description of Artificial Sequence;fusion protein  
 ; OTHER INFORMATION: of ECD and PD of human HER-2/neu  
 ; US-09-854-356-6

Query Match 19.8%; Score 83; DB 9; Length 919;  
 Best Local Similarity 100.0%; Pred. No. 7.5e-68; Mismatches 0; Indels 0; Gaps 0;  
 Matches 83; Conservative 0; APPLICANT: Stuart, Susan G.

QY 199 CKGSRCMGESEEDCQSLLRTVCAGGCARCKGGLPLTDCCHEOCAGCTGPKHSDCCLACIHF 258  
 Db 199 CKGSRCMGESEEDCQSLLRTVCAGGCARCKGGLPLTDCCHEOCAGCTGPKHSDCCLACIHF 258

QY 259 NHSGICELHCPALVNTDFES 281  
 Db 259 NHSGICELHCPALVNTDFES 281

RESULT 11  
 US-09-769-508-2  
 ; Sequence 2, Application US/09769508  
 ; Patent No. US20020155527A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Stuart, Susan G.  
 ; APPLICANT: Monahan, John J.  
 ; APPLICANT: Langton, Beatrice Claudia  
 ; APPLICANT: Hancock, Miriam E.C.  
 ; APPLICANT: Chao, Lorraine A.  
 ; APPLICANT: Bluford, Peper  
 ; TITLE OF INVENTION: C-ERBB-2 EXTERNAL DOMAIN: GP75  
 ; FILE REFERENCE: BE010-111-C1  
 ; CURRENT APPLICATION NUMBER: US/09/769,508-  
 ; CURRENT FILING DATE: 2001-01-26  
 ; NUMBER OF SEQ ID NOS: 2  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 2  
 ; LENGTH: 1255  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; US-09-769-508-2

RESULT 12  
 US-09-854-356-1  
 ; Sequence 1, Application US/09854356  
 ; Patent No. US20020177567A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Cheever, Martin A.  
 ; APPLICANT: Gheysen, Dirk  
 ; APPLICANT: Corixa Corporation  
 ; TITLE OF INVENTION: SmithKline Beecham Biologicals S. A.  
 ; FILE REFERENCE: 014058-009810PC  
 ; CURRENT APPLICATION NUMBER: US/09/854,356  
 ; CURRENT FILING DATE: 2001-05-09  
 ; PRIORITY APPLICATION NUMBER: US 09/493,480  
 ; PRIORITY FILING DATE: 2000-01-28  
 ; PRIORITY APPLICATION NUMBER: US 60/117,976  
 ; PRIORITY FILING DATE: 1999-01-29  
 ; NUMBER OF SEQ ID NOS: 26  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 1  
 ; LENGTH: 1255  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; OTHER INFORMATION: human HER-2/neu protein  
 ; NAME/KEY: DOMAIN  
 ; LOCATION: (1)..(653)  
 ; OTHER INFORMATION: extracellular domain (ECD)  
 ; NAME/KEY: DOMAIN  
 ; LOCATION: (676)..(1255)  
 ; OTHER INFORMATION: intracellular domain (ICD)  
 ; NAME/KEY: DOMAIN  
 ; LOCATION: (990)..(1255)  
 ; OTHER INFORMATION: phosphorylation domain (PD)  
 ; NAME/KEY: DOMAIN  
 ; LOCATION: (990)..(1048)  
 ; OTHER INFORMATION: fragment of the phosphorylation domain, preferred  
 ; OTHER INFORMATION: portion (delta PD)  
 ; US-09-854-356-1

Query Match 19.8%; Score 83; DB 9; Length 1255;  
 Best Local Similarity 100.0%; Pred. No. 9.8e-68; Mismatches 0; Indels 0; Gaps 0;  
 Matches 83; Conservative 0; APPLICANT: Stuart, Susan G.

QY 199 CKGSRCMGESEEDCQSLLRTVCAGGCARCKGGLPLTDCCHEOCAGCTGPKHSDCCLACIHF 258  
 Db 199 CKGSRCMGESEEDCQSLLRTVCAGGCARCKGGLPLTDCCHEOCAGCTGPKHSDCCLACIHF 258

QY 259 NHSGICELHCPALVNTDFES 281  
 Db 259 NHSGICELHCPALVNTDFES 281

RESULT 13  
 US-09-930-125-2  
 ; Sequence 2, Application US/09930125  
 ; Publication No. US2002019329A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Hand-Zimmerman, Susan  
 ; APPLICANT: Cheever, Martin A.  
 ; APPLICANT: Foy, Teresa M.  
 ; APPLICANT: Lodes, Michael J.  
 ; APPLICANT: Kalos, Michael D.

APPLICANT: McNeall, Patricia D.  
 TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND DIAGNOSIS  
 OF HER-2/NEU-ASSOCIATED MALIGNANCIES  
 FILE REFERENCE: 210121\_544  
 CURRENT APPLICATION NUMBER: US/09/930,125  
 CURRENT FILING DATE: 2001-08-14  
 NUMBER OF SEQ ID NOS: 25  
 SOFTWARE: FastSEQ for Windows Version 3.0  
 SEQ ID NO 2  
 LENGTH: 1255  
 TYPE: PRT  
 ORGANISM: Homo sapien  
 US-09-930-125-2

Query Match 19 8%; Score 83; DB 9; Length 1255;  
 Best Local Similarity 100.0%; Pred. No. 9.8e-68;  
 Matches 83; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Matches 83; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 199 CKGSRCGESSEDCQLSLTRIVCAGGCARCKGPLPTDCCHEQCAAGCTGPKHSCLACLF 258  
 Db 199 CKGSRCGESSEDCQLSLTRIVCAGGCARCKGPLPTDCCHEQCAAGCTGPKHSCLACLF 258  
 Qy 259 NHSGICELHCPALVNTNTFES 281  
 Db 259 NHSGICELHCPALVNTNTFES 281

RESULT 14  
 US-09-811-123-9  
 Sequence 9, Application US/09811123  
 Patent No. US2002001587A1  
 GENERAL INFORMATION:  
 APPLICANT: Sharon Erickson  
 APPLICANT: Ralph Schwall  
 APPLICANT: Mark Slinkowski  
 TITLE OF INVENTION: METHODS OF TREATMENT USING ANTI-ERBB  
 FILE REFERENCE: GENENT.073A2  
 CURRENT APPLICATION NUMBER: US/09/811,123  
 CURRENT FILING DATE: 2001-03-16  
 PRIOR APPLICATION NUMBER: 60/238,327  
 PRIOR FILING DATE: 2000-10-05  
 PRIOR APPLICATION NUMBER: 09/602,530  
 PRIOR FILING DATE: 2000-06-23  
 NUMBER OF SEQ ID NOS: 11  
 SOFTWARE: FastSEQ for Windows Version 4.0  
 SEQ ID NO 9  
 LENGTH: 1255  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 US-09-811-123-9

Query Match 19 8%; Score 83; DB 10; Length 1255;  
 Best Local Similarity 100.0%; Pred. No. 9.8e-68;  
 Matches 83; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Matches 83; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 199 CKGSRCGESSEDCQLSLTRIVCAGGCARCKGPLPTDCCHEQCAAGCTGPKHSCLACLF 258  
 Db 199 CKGSRCGESSEDCQLSLTRIVCAGGCARCKGPLPTDCCHEQCAAGCTGPKHSCLACLF 258  
 Qy 259 NHSGICELHCPALVNTNTFES 281  
 Db 259 NHSGICELHCPALVNTNTFES 281

Search completed: January 14, 2003, 17:21:51  
 Job time : 16.3032 secs

Query Match 19 8%; Score 83; DB 10; Length 1255;  
 Best Local Similarity 100.0%; Pred. No. 9.8e-68;  
 Matches 83; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Matches 83; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 199 CKGSRCGESSEDCQLSLTRIVCAGGCARCKGPLPTDCCHEQCAAGCTGPKHSCLACLF 258  
 Db 199 CKGSRCGESSEDCQLSLTRIVCAGGCARCKGPLPTDCCHEQCAAGCTGPKHSCLACLF 258  
 Qy 259 NHSGICELHCPALVNTNTFES 281  
 Db 259 NHSGICELHCPALVNTNTFES 281

RESULT 15  
 US-09-811-115-3  
 Sequence 3, Application US/09811115  
 Patent No. US20020035736A1  
 GENERAL INFORMATION:  
 APPLICANT: Erickson, Sharon  
 APPLICANT: Schwall, Ralph

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.3  
copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model!  
Run on: January 14, 2003, 17:11:49 ; Search time 26.9237 Seconds  
(without alignments)  
457.894 Million cell updates/sec

Title: US-09-506-079E-2  
Perfect score: 419  
Sequence: 1 MELALALCRWGLLALLLPPGA.....VGRGXDPDAHAVAVLSRYEG 419

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 262574 seqs, 29422922 residues

Word size : 0

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : Issued\_Patents\_AN:\*

1: /cggn2\_6/ptodata/1/1aa/5A.COMB.pep:\*

2: /cggn2\_6/ptodata/1/1aa/5B.COMB.pep:\*

3: /cggn2\_6/ptodata/1/1aa/6A.COMB.pep:\*

4: /cggn2\_6/ptodata/1/1aa/6B.COMB.pep:\*

5: /cggn2\_6/ptodata/1/1aa/backfilest1.pep:\*

Pred.: No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query	Length	DB ID	Description
1	83	19.8	419	4	US-09-630-155-2
2	83	19.8	624	3	US-08-422-108-1
3	83	19.8	624	4	US-08-422-734-1
4	83	19.8	782	2	US-08-348A-4
5	83	19.8	782	3	US-08-579-823A-4
6	83	19.8	782	4	US-09-344-195-4
7	83	19.8	1255	1	US-08-467-083-68
8	83	19.8	1255	1	US-08-414-417B-68
9	83	19.8	1255	2	US-08-484-438B-8
10	83	19.8	1255	2	US-08-348A-68
11	83	19.8	1255	2	US-08-625-101-2
12	83	19.8	1255	2	US-08-468-545B-68
13	83	19.8	1255	3	US-08-356-786-2
14	83	19.8	1255	3	US-08-466-680B-68
15	52	12.4	97	1	US-08-421-356-3
16	17	4.1	17	1	US-08-467-083-3
17	17	4.1	17	1	US-08-467-083-61
18	17	4.1	17	1	US-08-414-417B-61
19	17	4.1	17	2	US-08-468-545B-61
20	17	4.1	17	2	US-08-468-545B-61
21	17	4.1	17	3	US-08-466-680B-61
22	15	3.6	15	1	US-08-467-083-30
23	15	3.6	15	1	US-08-467-083-31
24	15	3.6	15	1	US-08-467-083-56
25	15	3.6	15	1	US-08-414-417B-30
26	15	3.6	15	1	US-08-414-417B-31
27	15	3.6	15	1	US-08-414-417B-56

#### ALIGNMENTS

RESULT 1  
US-09-630-155-2  
Sequence 2, Application US/9630155  
; Patent No. 6414130

GENERAL INFORMATION:

APPLICANT: Doherty, Joni Kristin and Gail M. Clinton  
TITLE OF INVENTION: HER-2 BINDING ANTAGONISTS

NUMBER OF SEQUENCES: 9

CORRESPONDENCE ADDRESS:  
NAME: Davidson, Barry L.  
REGISTRATION NUMBER: 47,309  
REFERENCE/DOCKET NUMBER: 49321-10

TELECOMMUNICATION INFORMATION:  
STREET: 1501 Fourth Avenue, 2600 Century Square  
CITY: Seattle  
STATE: Washington  
COUNTRY: U.S.A.  
ZIP: 98101

SOFTWARE: Word  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/630,155  
FILING DATE: 16-Jan-2001  
CLASSIFICATION: <Unknown>

ATTORNEY/AGENT INFORMATION:  
NAME: Davidson, Barry L.  
REGISTRATION NUMBER: 47,309  
REFERENCE/DOCKET NUMBER: 49321-10

TELEPHONE: 206 628-7621  
TELEFAX: 206 628-7699

INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 419  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: unknown  
MOLECULE TYPE: polypeptide  
SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
US-09-630-155-2

Query Match Best Local Similarity 19.8%; Score 83; DB 4; Length 419;  
Matches 83; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 199 CKGSRCGWGESSEDQSLTRIVCAGGCARKGKPLPTDCCLBQCAAGCTGPRHSDCLACLIR 258  
Db 199 CKGSRCGWGESSEDQSLTRIVCAGGCARKGKPLPTDCCLBQCAAGCTGPRHSDCLACLIR 258

OY 259 NHSGICELHC PALVNTDFFES 281

RESULT 2  
US-08-422-108-1  
Sequence 1, Application US/08422108  
; Patent No. 6015567  
; GENERAL INFORMATION:  
; APPLICANT: Hudziak, Robert M.  
; APPLICANT: Ulrich, Axel  
TITLE OF INVENTION: HER2 EXTRACELLULAR DOMAIN  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genentech, Inc.  
; STREET: 460 Point San Bruno Blvd  
; CITY: South San Francisco  
; STATE: California  
; ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WinPatin (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/422,108  
FILING DATE: 14-APR-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/355460  
FILING DATE: 13-DEC-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/422,108  
FILING DATE: 14-APR-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/354319  
FILING DATE: 19-MAY-1989  
ATTORNEY/AGENT INFORMATION:  
NAME: Lee, Wendy M.  
REGISTRATION NUMBER: 00, 000  
REFERENCE/DOCKET NUMBER: 554C2D2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415/225-9881  
TELEFAX: 415/952-9881  
TELEX: 910/371-7168  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 624 amino acids  
TYPE: Amino Acid  
TOPOLOGY: Linear  
US-08-422-108-1  
Query Match 19.8%; Score 83; DB 3; Length 624;  
Best Local Similarity 100.0%; Pred. No. 2.2e-71; Mismatches 0; Indels 0; Gaps 0;  
Matches 83; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Oy 199 CKGSRQWGESESSDCOSLRTWCAGGCCARCKPLPTCOCHEOCAAGCTGPKHSDCLACLHF 258  
Db 178 CKGSRQWGESESSDCOSLRTWCAGGCCARCKPLPTCOCHEOCAAGCTGPKHSDCLACLHF 237  
Qy 259 NHSGICELHCPALVTNTDIFES 281  
Db 238 NHSGICELHCPALVTNTDIFES 260  
RESULT 4  
US-09-146-283-4  
Sequence 4, Application US/09146283  
; Patent No. 5916546  
; GENERAL INFORMATION:  
; APPLICANT: Laus, Reiner  
; APPLICANT: Ruegg, Curtiss L.  
; APPLICANT: Wu, Hongyu  
TITLE OF INVENTION: Immunostimulatory Compositions  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Dehlinger & Associates  
; STREET: 350 Cambridge Ave. Suite 250  
; CITY: Palo Alto  
RESULT 3  
US-08-422-734-1  
Sequence 1, Application US/08422734  
; Patent No. 6333169  
; GENERAL INFORMATION:  
; APPLICANT: Hudziak, Robert M.

STATE: CA  
COUNTRY: USA  
ZIP: 94106

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08-579, 823A  
FILING DATE: 03-DEC-1998  
CLASSIFICATION: 536

ATTORNEY/AGENT INFORMATION:  
NAME: Judge, Linda R.  
REGISTRATION NUMBER: 42,702

REFERENCE/DOCKET NUMBER: 7636-0010  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650-324-0880  
TELEFAX: 650-324-0960

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:  
LENGTH: 782 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ORIGINAL SOURCE:  
ORGANISM: homo sapiens  
INDIVIDUAL ISOLATE: GM-CSF-Her-2 fusion protein; Fig. 8  
9-146-283-4

RESULT 5  
US-08-579-823A-4

	Query Match	Score	DB	Length
Matches	Best Local Similarity	83;	2;	782;
Db	Patent No.	0;	2,76-71;	Mismatches
Qy	GENERAL INFORMATION:	0;	Indels	0;
Qy	APPLICANT: Laus, Reiner	0;	Gaps	0;
Qy	TITLE OF INVENTION: Immunostimulatory Compositions	0;		
Qy	NUMBER OF SEQUENCES: 10	0;		
Qy	CORRESPONDENCE ADDRESS:	0;		
Qy	ADDRESSEE: Behlinder & Associates	0;		
Qy	STREET: 350 Cambridge Ave. Suite 250	0;		
Qy	CITY: Palo Alto	0;		
Qy	STATE: CA	0;		
Qy	COUNTRY: USA	0;		

RESULT 6  
US-09-344-195-4

	Query Match	Score	DB	Length
Matches	Best Local Similarity	83;	3;	782;
Db	Patent No.	0;	2,76-71;	Mismatches
Qy	GENERAL INFORMATION:	0;	Indels	0;
Qy	APPLICANT: Laus, Reiner	0;	Gaps	0;
Qy	TITLE OF INVENTION: Immunostimulatory Compositions	0;		
Qy	NUMBER OF SEQUENCES: 10	0;		
Qy	CORRESPONDENCE ADDRESS:	0;		
Qy	ADDRESSEE: Behlinder & Associates	0;		
Qy	STREET: 350 Cambridge Ave. Suite 250	0;		
Qy	CITY: Palo Alto	0;		
Qy	STATE: CA	0;		
Qy	COUNTRY: USA	0;		

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/344, 195  
FILING DATE: 24-Jun-1999  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/09/146, 283  
FILING DATE: 03-SEPT-1998  
ATTORNEY/AGENT INFORMATION:  
NAME: Judge, Linda R.  
REGISTRATION NUMBER: 42,702

REFERENCE/DOCKET NUMBER: 7636-0010.21  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650-324-0880  
TELEFAX: 650-324-0960

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:  
LENGTH: 782 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ORIGINAL SOURCE:  
ORGANISM: homo sapiens  
INDIVIDUAL ISOLATE: GM-CSF-Her-2 fusion protein; Fig. 8  
US-08-579-823A-4

ORGANISM: homo sapiens  
 INDIVIDUAL ISOLATE: GM-CSF-Her-2 fusion protein; FIG. 8  
 SEQUENCE DESCRIPTION: SEQ ID NO: 4;  
 US-09-344-195-4

RESULT 7  
 US-08-467-083-68  
 Sequence 68, Application US/08467083  
 Patent No. 5726023  
 GENERAL INFORMATION:  
 APPLICANT: Cheever, Martin A.  
 APPLICANT: Disis, Mary L.  
 TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/neu PROTEIN FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE TITLE OF INVENTION: FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE NUMBER OF SEQUENCES: 69  
 NUMBER OF INVENTIONS: HER-2/neu ONCOGENE IS ASSOCIATED  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Seed and Berry LLP  
 STREET: 6300 Columbia Center, 701 Fifth Avenue  
 CITY: Seattle  
 STATE: Washington  
 ZIP: 98104-7092  
 COMPUTER READABLE FORM:  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent In Release #1.0, version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/414,417B  
 FILING DATE: 31-MAR-1995  
 CLASSIFICATION: 424  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Sharkey, Richard G.  
 REGISTRATION NUMBER: 32,629  
 REFERENCE/DOCKET NUMBER: 920010.44BC2  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (206) 682-4900  
 TELEFAX: (206) 682-6031  
 INFORMATION FOR SEQ ID NO: 68:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1255 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear

RESULT 8  
 US-08-414-417B-68  
 Sequence 68, Application US/08414417B  
 Patent No. 5801005  
 GENERAL INFORMATION:  
 APPLICANT: Cheever, Martin A.  
 APPLICANT: Disis, Mary L.  
 TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/neu PROTEIN FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE TITLE OF INVENTION: FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE NUMBER OF SEQUENCES: 69  
 NUMBER OF INVENTIONS: HER-2/neu ONCOGENE IS ASSOCIATED  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Seed and Berry LLP  
 STREET: 6300 Columbia Center, 701 Fifth Avenue  
 CITY: Seattle  
 STATE: Washington  
 ZIP: 98104-7092  
 COMPUTER READABLE FORM:  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent In Release #1.0, version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/414,417B  
 FILING DATE: 31-MAR-1995  
 CLASSIFICATION: 424  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Sharkey, Richard G.  
 REGISTRATION NUMBER: 32,629  
 REFERENCE/DOCKET NUMBER: 920010.44BC2  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (206) 682-4900  
 TELEFAX: (206) 682-6031  
 INFORMATION FOR SEQ ID NO: 68:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1255 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear

RESULT 9  
 US-08-484-438-8  
 Sequence 8, Application US/08484438  
 Patent No. 5811098 5780031  
 GENERAL INFORMATION:  
 APPLICANT: Plowman, Gregory D.  
 APPLICANT: Culouscou, Jean-Michel  
 APPLICANT: Shovar, Mohammed  
 APPLICANT: Siegall, Clay B.  
 APPLICANT: Hellstr m, Ingemar  
 APPLICANT: Hellstr m, Karl E.  
 TITLE OF INVENTION: HER4 HUMAN RECEPTOR TYROSINE KINASE  
 NUMBER OF SEQUENCES: 42  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Pennie & Edmonds

RESULT 7  
 US-08-467-083-68  
 Sequence 68, Application US/08467083  
 Patent No. 5726023  
 GENERAL INFORMATION:  
 APPLICANT: Cheever, Martin A.  
 APPLICANT: Disis, Mary L.  
 TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/neu PROTEIN FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE TITLE OF INVENTION: FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE NUMBER OF SEQUENCES: 69  
 NUMBER OF INVENTIONS: HER-2/neu ONCOGENE IS ASSOCIATED  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Seed and Berry LLP  
 STREET: 6300 Columbia Center, 701 Fifth Avenue  
 CITY: Seattle  
 STATE: Washington  
 ZIP: 98104-7092  
 COMPUTER READABLE FORM:  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent In Release #1.0, version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/414,417B  
 FILING DATE: 31-MAR-1995  
 CLASSIFICATION: 424  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Sharkey, Richard G.  
 REGISTRATION NUMBER: 32,629  
 REFERENCE/DOCKET NUMBER: 920010.44BC2  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (206) 682-4900  
 TELEFAX: (206) 682-6031  
 INFORMATION FOR SEQ ID NO: 68:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1255 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear

RESULT 8  
 US-08-414-417B-68  
 Sequence 68, Application US/08414417B  
 Patent No. 5801005  
 GENERAL INFORMATION:  
 APPLICANT: Cheever, Martin A.  
 APPLICANT: Disis, Mary L.  
 TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/neu PROTEIN FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE TITLE OF INVENTION: FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE NUMBER OF SEQUENCES: 69  
 NUMBER OF INVENTIONS: HER-2/neu ONCOGENE IS ASSOCIATED  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Seed and Berry LLP  
 STREET: 6300 Columbia Center, 701 Fifth Avenue  
 CITY: Seattle  
 STATE: Washington  
 ZIP: 98104-7092  
 COMPUTER READABLE FORM:  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent In Release #1.0, version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/414,417B  
 FILING DATE: 31-MAR-1995  
 CLASSIFICATION: 424  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Sharkey, Richard G.  
 REGISTRATION NUMBER: 32,629  
 REFERENCE/DOCKET NUMBER: 920010.44BC2  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (206) 682-4900  
 TELEFAX: (206) 682-6031  
 INFORMATION FOR SEQ ID NO: 68:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1255 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear

RESULT 9  
 US-08-484-438-8  
 Sequence 8, Application US/08484438  
 Patent No. 5811098 5780031  
 GENERAL INFORMATION:  
 APPLICANT: Plowman, Gregory D.  
 APPLICANT: Culouscou, Jean-Michel  
 APPLICANT: Shovar, Mohammed  
 APPLICANT: Siegall, Clay B.  
 APPLICANT: Hellstr m, Ingemar  
 APPLICANT: Hellstr m, Karl E.  
 TITLE OF INVENTION: HER4 HUMAN RECEPTOR TYROSINE KINASE  
 NUMBER OF SEQUENCES: 42  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Pennie & Edmonds

Query Match 19.8%; Score 83; DB 1; Length 1255;  
 Best Local Similarity 100.0%; Pred. No. 4.2e-71;  
 Matches 83; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 199 CKGSRGCGESSEDCQSITRTVAGGCCARKGKGPLPTDCHEQCAAGCTGPKHSCLACLHF 258  
 Db 199 CKGSRGCGESSEDCQSITRTVAGGCCARKGKGPLPTDCHEQCAAGCTGPKHSCLACLHF 258  
 Qy 259 NHSGICELHCPALVNTDFFS 281  
 |||||||  
 Query Match 19.8%; Score 83; DB 1; Length 1255;  
 Best Local Similarity 100.0%; Pred. No. 4.2e-71;  
 Matches 83; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 199 CKGSRGCGESSEDCQSITRTVAGGCCARKGKGPLPTDCHEQCAAGCTGPKHSCLACLHF 258  
 Db 199 CKGSRGCGESSEDCQSITRTVAGGCCARKGKGPLPTDCHEQCAAGCTGPKHSCLACLHF 258  
 Qy 259 NHSGICELHCPALVNTDFFS 281  
 |||||||

STREET: 1155 Avenue of the Americas  
 CITY: New York  
 STATE: New York  
 COUNTRY: U.S.A.  
 ZIP: 10036-2711  
 COMPUTER READABLE FORM:  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/484,438  
 FILING DATE: 07-JUN-1995  
 CLASSIFICATION: 530  
 PRIORITY APPLICATION DATA:  
 APPLICATION NUMBER: 08/323,442  
 FILING DATE: 14-OCT-1994  
 APPLICATION NUMBER: US 08/150,704  
 FILING DATE: 10-NOV-1993  
 CLASSIFICATION: 530  
 PRIORITY APPLICATION DATA:  
 APPLICATION NUMBER: US 07/981,165  
 FILING DATE: 24-NOV-1992  
 CLASSIFICATION: 530  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Mistock, S. Leslie  
 REGISTRATION NUMBER: 18,872  
 REFERENCE/DOCKET NUMBER: 5624-230  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (212) 790-9090  
 TELEFAX: (212) 869-8864/9741  
 INFORMATION FOR SEQ ID NO: 8:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1255 amino acids  
 TYPE: amino acid  
 STRANDBNESS: unknown  
 TOPOLogy: unknown  
 MOLECULE TYPE: protein  
 US-08-484-438-8

Query Match 19.8%; Score 83; DB 2; length 1255;  
 Best Local Similarity 100.0%; Pred. No. 4.2e-71;  
 Matches 83; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 199 CKGSRCGWGSSEDCQSLTRIVCAGGCARCKGPLPTDCHEQCAAGCTGPKHSCLACIHF 258  
 Db 199 CKGSRCGWGSSEDCQSLTRIVCAGGCARCKGPLPTDCHEQCAAGCTGPKHSCLACIHF 258  
 QY 259 NHSGICELHCPALVNTDFES 281  
 Db 259 NHSGICELHCPALVNTDFES 281

RESULT 11  
 US-08-625-101-2  
 Sequence 2: Application US/08625101  
 Patent No. 5869445  
 GENERAL INFORMATION:  
 APPLICANT: Cheever, Martin A.  
 APPLICANT: Disis, Mary L.  
 TITLE OF INVENTION: COMPOUNDS FOR ELICITING OR ENHANCING IMMUNE REACTIVITY TO HER-2/neu PROTEIN FOR PREVENTION OF MALIGNANCIES IN WHICH THE HER-2/neu TREATMENT OF MALIGNANCIES OR  
 TITLE OF INVENTION: OR TREATMENT OF MALIGNANCIES IN WHICH THE HER-2/neu NUMBER OF SEQUENCES: 4  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: SEED and BERRY LLP  
 STREET: 6300 Columbia Center, 701 Fifth Avenue  
 CITY: Seattle  
 STATE: Washington  
 COUNTRY: USA  
 ZIP: 98104-7092

COMPUTER READABLE FORM:  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/625,101  
 FILING DATE: 01-APR-1996  
 CLASSIFICATION: 424  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Sharkey, Richard G.  
 REGISTRATION NUMBER: 32,629  
 REFERENCE/DOCKET NUMBER: 920010.448C7  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (206) 622-4900  
 TELEFAX: (206) 682-6031  
 INFORMATION FOR SEQ ID NO: 2:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1255 amino acids

US-08-484-438-8

RESULT 10  
 US-08-480-348A-68  
 Sequence 68: Application US/084806348A  
 Patient No. 5846538  
 GENERAL INFORMATION:  
 APPLICANT: Cheever, Martin A.  
 APPLICANT: Disis, Mary L.  
 TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/neu PROTEIN FOR DIAGNOSTS AND TREATMENT OF MALIGNANCIES IN WHICH THE HER-2/neu  
 TITLE OF INVENTION: HER-2/neu ONCOGENE IS ASSOCIATED  
 NUMBER OF SEQUENCES: 69  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Seed and Berry LLP  
 STREET: 6300 Columbia Center, 701 Fifth Avenue  
 CITY: Seattle  
 STATE: Washington  
 COUNTRY: USA  
 ZIP: 98104-7092

COMPUTER READABLE FORM:  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/625,101  
 FILING DATE: 01-APR-1996  
 CLASSIFICATION: 424  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Sharkey, Richard G.  
 REGISTRATION NUMBER: 32,629  
 REFERENCE/DOCKET NUMBER: 920010.448C7  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (206) 622-4900  
 TELEFAX: (206) 682-6031  
 INFORMATION FOR SEQ ID NO: 2:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1255 amino acids

US-08-480-348A-68

; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-625-101-2

US-08-356-786-2  
; Sequence 2, Application US/08356786  
; Patent No. 5877305  
; GENERAL INFORMATION:  
; APPLICANT: Huston, James S.  
; APPLICANT: Opfermann, Hermann  
; APPLICANT: Houston, L. L.  
; APPLICANT: King, David B.

Query Match 19.8%; Score 83; DB 2; Length 1255;  
Best Local Similarity 100.0%; Pred. No. 4.2e-71; Mismatches 0; Indels 0; Gaps 0;  
Matches 83; Conservative 0; Mismatches 0; Gaps 0;

Qy 199 CKGSRGESSEDCQLSLRTVCAGGGCARCKGPLPTDCHEOCAGCTGPKHSCLACLF 258  
Db 199 CKGSRGESSEDCQLSLRTVCAGGGCARCKGPLPTDCHEOCAGCTGPKHSCLACLF 258

Qy 259 NHSGICELHCPALVNTDFFES 281  
Db 259 NHSGICELHCPALVNTDFFES 281

RESULT 12  
US-08-468-545B-68  
; Sequence 68, Application US/08468545B  
; Patent No. 5876712  
; GENERAL INFORMATION:  
; APPLICANT: Cheever, Martin A.  
; APPLICANT: Disis, Mary L.  
; TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/neu PROTEIN  
; TITLE OF INVENTION: FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE  
; TITLE OF INVENTION: HER-2/neu ONCOGENE IS ASSOCIATED  
; NUMBER OF SEQUENCES: 69  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Seed and Berry LLP  
; STREET: 6300 Columbia Center, 701 Fifth Avenue  
; CITY: Seattle  
; STATE: Washington  
; COUNTRY: US  
; ZIP: 98104-7092  
; COMPUTER READABLE FORM:  
; COMPUTER TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patientin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/356,786  
; FILING DATE: 06-FEB-1992  
; PRIORITY APPLICATION DATA:  
; APPLICATION NUMBER: 07/831,967  
; FILING DATE: 06-FEB-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Pitcher, Edmund R.  
; REGISTRATION NUMBER: 27,829  
; REFERENCE/DOCKET NUMBER: CRP-053  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 248-7100  
; TELEFAX: (617) 248-7100  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1255 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-356-786-2

Query Match 19.8%; Score 83; DB 2; Length 1255;  
Best Local Similarity 100.0%; Pred. No. 4.2e-71; Mismatches 0; Indels 0; Gaps 0;  
Matches 83; Conservative 0; Mismatches 0; Gaps 0;

Qy 199 CKGSRGESSEDCQLSLRTVCAGGGCARCKGPLPTDCHEOCAGCTGPKHSCLACLF 258  
Db 199 CKGSRGESSEDCQLSLRTVCAGGGCARCKGPLPTDCHEOCAGCTGPKHSCLACLF 258

Qy 259 NHSGICELHCPALVNTDFFES 281  
Db 259 NHSGICELHCPALVNTDFFES 281

RESULT 14  
US-08-466-680B-68  
; Sequence 68, Application US/08466680B  
; Patent No. 6075122  
; GENERAL INFORMATION:  
; APPLICANT: Cheever, Martin A.  
; APPLICANT: Disis, Mary L.  
; TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/neu PROTEIN  
; TITLE OF INVENTION: FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE  
; NUMBER OF SEQUENCES: 69  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Seed and Berry LLP  
; STREET: 6300 Columbia Center, 701 Fifth Avenue  
; CITY: Seattle

RESULT 13



**THIS PAGE BLANK (USPTO)**

Gencore version 5.1.3  
copyright (c) 1993 - 2003 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: January 14, 2003, 17:08:59 ; Search time 53.8474 Seconds  
 Sequence: 1 MELANICRWGLLALLLPPGA.....VGRGXDPDAHVAVXLSREG 419  
 (without alignments)  
 1036.856 Million cell updates/sec

## Title: US-09-506-079E-2

Perfect score: 419  
 Sequence: 1 MELANICRWGLLALLLPPGA.....VGRGXDPDAHVAVXLSREG 419

Scoring table: OLIGO Gapop 60.0 , Gapext 60.0

Searched: 908470 seqs, 133250620 residues

Word size : 0

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0  
 Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : A\_Geneseq\_101002.\*  
 1: /SIDS2/gcadata/geneseq/geneseq -emb1/AA1980.DAT: \*  
 2: /SIDS2/gcadata/geneseq/geneseq -emb1/AA1981.DAT: \*  
 3: /SIDS2/gcadata/geneseq/geneseq -emb1/AA1982.DAT: \*  
 4: /SIDS2/gcadata/geneseq/geneseq -emb1/AA1984.DAT: \*  
 5: /SIDS2/gcadata/geneseq/geneseq -emb1/AA1985.DAT: \*  
 6: /SIDS2/gcadata/geneseq/geneseq -emb1/AA1987.DAT: \*  
 7: /SIDS2/gcadata/geneseq/geneseq -emb1/AA1989.DAT: \*  
 9: /SIDS2/gcadata/geneseq/geneseq -emb1/AA1991.DAT: \*  
 10: /SIDS2/gcadata/geneseq/geneseq -emb1/AA1992.DAT: \*  
 11: /SIDS2/gcadata/geneseq/geneseq -emb1/AA1993.DAT: \*  
 12: /SIDS2/gcadata/geneseq/geneseq -emb1/AA1994.DAT: \*  
 13: /SIDS2/gcadata/geneseq/geneseq -emb1/AA1995.DAT: \*  
 14: /SIDS2/gcadata/geneseq/geneseq -emb1/AA1996.DAT: \*  
 15: /SIDS2/gcadata/geneseq/geneseq -emb1/AA1997.DAT: \*  
 16: /SIDS2/gcadata/geneseq/geneseq -emb1/AA1998.DAT: \*  
 17: /SIDS2/gcadata/geneseq/geneseq -emb1/AA2000.DAT: \*  
 18: /SIDS2/gcadata/geneseq/geneseq -emb1/AA2001.DAT: \*  
 19: /SIDS2/gcadata/geneseq/geneseq -emb1/AA2002.DAT: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARES

Result No.	Query	% Match	Length	DB ID	Description
1	284	67.8	419	22 AAE09181	Human p68HER-2 gen.
2	284	67.8	419	22 AAE09208	Human p68HER-2 gen.
3	284	67.8	419	23 AAE0348	Human truncated HE
4	278	66.3	419	22 AAE05212	Human p68HER-2 gen.
5	278	66.3	419	22 AAE09211	Human p68HER-2 gen.
6	261	62.3	419	22 AAE09210	Human p68HER-2 gen.
7	244	58.2	419	22 AAE09209	Human p68HER-2 gen.
8	239	57.0	419	22 AAE09202	Human p68HER-2 gen.
9	228	54.4	419	22 AAE09207	Human p68HER-2 gen.
10	227	54.2	419	22 AAE09183	Human p68HER-2 gen.

## ALIGNMENTS

RESULT 1  
 ID AAE09181  
 AAED9181 standard; Protein: 419 AA.

ID XX  
 AAED9181;  
 AC XX  
 DT 15-NOV-2001 (first entry)

DE Human p68HER-2 generic sequence #1.  
 XX HER-2; herstatin; antagonist; extracellular domain; ECD; Herceptin;  
 KW solid tumour; cancer; polymorphism; cytostatic; gene therapy;  
 KW p68HER-2; ECDIIA.

XX Homo sapiens.

OS Location/Qualifiers  
 XX 1..340  
 FH /note= "Identical to N-terminal region of p185HER-2"  
 FT Key  
 FT Region  
 FT Domain  
 FT Label= ECDIIA  
 FT /note= "Extracellular domain IIIa".

FT Misc-difference 124  
 FT /note= "Represented as Arg in the sequence shown in the specification".

FT Misc-difference 125  
 FT /note= "Represented as Arg in the sequence shown in the specification".

FT Misc-difference 342  
 FT /label= Unknown  
 FT /note= "Encoded by WCC".

FT Misc-difference 345  
 FT /label= Unknown

FT Misc-difference /note= "Encoded by CYG" 346  
 FT /label= Unknown 346  
 FT /note= "Encoded by CYC" 356  
 FT Misc-difference /label= Unknown 356  
 FT /note= "Encoded by CWG" 358  
 FT /label= Unknown 358  
 FT /note= "Encoded by ATR" 361  
 FT Misc-difference /label= Unknown 361  
 FT /note= "Encoded by GNC" 366  
 FT /label= Unknown 366  
 FT /note= "Encoded by MTA" 369  
 FT Misc-difference /note= "Encoded by AGC" 394  
 FT Misc-difference /label= Unknown 394  
 FT /note= "Encoded by CST" 404  
 FT Misc-difference /label= Unknown 404  
 FT /note= "Encoded by CYG" 413  
 FT Misc-difference /label= Unknown 413  
 FT /note= "Encoded by SAC" 413  
 PN WO200161356-A1.  
 XX  
 PD 23-AUG-2001.  
 PR 16-FEB-2001; 2001WO-US05327.  
 PR 16-FEB-2000; 2000US-0506079.  
 PR (UYOR ) UNIV OREGON HEALTH SCI.  
 PT Clinton G, Henner WD, Evans A;  
 XX  
 DR WPI; 2001-529934/58.  
 DR N-PSDB; AAD15844.  
 XX  
 PT New polypeptide, which binds to the extracellular domain of HER-2 for  
 the treatment of hard tumors -  
 XX  
 PS Claim 8; Page 53-54; 61PP; English.  
 XX  
 CC The invention relates to novel HER-2 (herstatin-2) antagonist  
 CC particularly a polypeptide that binds to the extracellular domain (ECO)  
 CC of HER-2 at a site that is different from the binding site of humanised  
 CC antibody, Herceptin, at an affinity of at least 10<sup>-8</sup>. The present  
 CC invention is based upon the initial discovery of an alternative HER-2  
 CC mRNA transcript with 274 bp insert of intron 8. The translation product  
 CC of the alternative transcript is a truncated HER-2 protein designated  
 CC p68HER-2 which lacks the transmembrane and intracellular domains of  
 CC p185HER-2 but contains ECD I, II of the p185HER-2 and the novel ECDIIA.  
 CC The ECDIIA-containing polypeptides bind tightly to, and thus antagonise  
 CC the HER-2 receptor. The peptides, which bind to an HER-2 ECD, and the  
 CC nucleic acids encoding these are useful to treat, diagnose and identify  
 CC solid tumours. The present sequence is human p68HER-2 protein  
 CC containing ECDIIA generic sequence.  
 XX  
 SQ Sequence 419 AA;

Query Match 67.8%; score 284; DB 22; Length 419;  
 Best Local Similarity 100.0%; Pred. No. 3.2e-259; Length 419;  
 Matches 294; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 126 TTPVIGASPGGIRELQLSLTECLKGKVLIQRNPQLCYDQMLILWWDIFHKNNQLAITLID 185  
 Db 126 TTPVIGASPGGIRELQLSLTECLKGKVLIQRNPQLCYDQMLILWWDIFHKNNQLAITLID 185

RESULT 2  
 AAE09208  
 ID AAE09208 standard; Protein: 419 AA.  
 XX  
 AC AAE09208;  
 XX  
 DT 15-NOV-2001 (first entry)  
 XX  
 DE Human p68HER-2 generic protein variant 6.  
 KW HER-2; herstatin; antagonist; extracellular domain; ECO; Herceptin;  
 KW p68HER-2; ECDIIA; variant.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Region 1..340  
 FT /note= "Identical to N-terminal region of p185HER-2"  
 FT Domain 341..419  
 FT /label= ECDIIA variant  
 FT /note= "Extracellular domain IIIA variant"  
 FT Misc-difference 124  
 FT /note= "Represented as Agn in the parent sequence shown  
 in the specification"  
 FT Misc-difference 125  
 FT /note= "Represented as Agn in the parent sequence shown  
 in the specification"  
 FT Misc-difference 342  
 FT /label= Unknown  
 FT Misc-difference 345  
 FT /label= Unknown  
 FT Misc-difference 346  
 FT /label= Unknown  
 FT Misc-difference 356  
 FT /label= Unknown  
 FT Misc-difference 358  
 FT /label= Unknown  
 FT Misc-difference 361  
 FT /label= Asp, Ala, Val  
 FT Misc-difference 376  
 FT /label= Unknown  
 FT Misc-difference 394  
 FT /label= Unknown  
 FT Misc-difference 404  
 FT /label= Unknown  
 FT Misc-difference 413  
 FT /label= Unknown  
 XX  
 PN WO200161356-A1.  
 XX  
 PD 23-AUG-2001.  
 XX  
 PR 16-FEB-2001; 2001WO-US05327.  
 PR 16-FEB-2000; 2000US-0506079.



306 TDVGSCTLVCPLNQEVTAEDGTORCEKSKPCARGXHSXXPRPAAVPVPPXRQXPXPAHP 365  
 PT PT New polypeptide, which binds to the extracellular domain of HER-2 for  
 Db 306 TDVGSCTLVCPLNQEVTAEDGTORCEKSKPCARGXHSXXPRPAAVPVPPXRQXPXPAHP 365  
 PT PT the treatment of hard tumors

Example 11; Page -, 61pp; English.

RESULT 4  
 AAE0912 ID AAE0912 standard; Protein; 419 AA.  
 AAE0912; XX  
 15-NOV-2001 (first entry)

XX Human p68HER-2 generic protein variant 10.

DE Human p68HER-2 generic protein variant 10.

KW HER-2; herstatin; antagonist; extracellular domain; ECD; Herceptin;  
 solid tumour; cancer; polymorphism; cytostatic; gene therapy;

KW p68HER-2; ECDIIIA; variant.

XX OS Homo sapiens.

XX Key Location/Qualifiers

FH Region 1..340

FT /note= "Identical to N-terminal region of p185HER-2"

FT Domain 341..419

FT /label= ECIIIA variant

FT /note= "Extracellular domain IIIA variant"

FT MISC-difference 124

FT /note= "Represented as Agn in the parent sequence shown in the specification"

FT MISC-difference 125

FT /note= "Represented as Agn in the parent sequence shown in the specification"

FT MISC-difference 342

FT /label= Unknown

FT MISC-difference 345

FT /label= Unknown

FT MISC-difference 346

FT /label= Unknown

FT MISC-difference 356

FT /label= Unknown

FT MISC-difference 358

FT /label= Unknown

FT MISC-difference 361

FT /label= Unknown

FT MISC-difference 376

FT /label= Unknown

FT MISC-difference 394

FT /label= Unknown

FT MISC-difference 404

FT /label= Unknown

FT MISC-difference 413

FT /note= "p68HER-2 generic sequence (AAE09181) xaa substituted with ASN"

XX WO200161356-A1.

XX PD 23-AUG-2001.

XX FT 16-FEB-2001; 2000US-0506079.

XX PR 16-FEB-2000; 2000US-0506079.

XX PA (UYOR-) UNIV OREGON HEALTH SCI.

XX Clinton G, Henner WD, Evans A;

XX PR WPI; 2001-529934/58.

XX Key Location/Qualifiers

FT Region 1..340

FT /note= "Identical to N-terminal region of p185HER-2"

FT Domain 341..419

FT /label= ECIIIA variant

FT /note= "Extracellular domain IIIA variant"

XX CC The invention relates to novel HER-2 (herstatin-2) antagonist particularly a polypeptide that binds to the extracellular domain (ECD) of HER-2 at a site that is different from the binding site of humanised antibody, Herceptin, at an affinity of at least 10<sup>8</sup>. The present invention is based upon the initial discovery of an alternative HER-2 mRNA transcript with 274 bp insert of intron 8. The translation product of the alternative transcript is a truncated HER-2 protein designated p68HER-2 which lacks the transmembrane and intracellular domains of p185HER-2 but contains ECD 1, 11 of the p185HER-2 and the novel ECIIIA. The ECIIIA-containing polypeptides bind tightly to, and thus antagonise the HER-2 receptor. The peptides, which bind to an HER-2 ECD, and the nucleic acids encoding these are useful to treat, diagnose and identify solid tumours. The present sequence is a truncated HER-2 generic protein containing ECIIIA variant sequence.

CC Note: The present sequence is not shown in the specification but is derived from p68HER-2 generic sequence (SEQ ID NO:2) shown in the sequence listing (AAE09181).

XX SQ Sequence 419 AA;

Query	Match	Score	DB	Length
Best Local Similarity	100.0%	278;	22;	419;
Matches	287;	Conservative	0;	Mismatches 0;
			Indels 0;	Gaps 0;

XX OY 126 TTPVTGASPGGLERLQRSLTICKLGVLQIQRNQLCYQDTIWKDFHKNNKLALTIID 185

Db 126 TTPVTGASPGGLERLQRSLTICKLGVLQIQRNQLCYQDTIWKDFHKNNKLALTIID 185

OY 186 TNRSRACHCPCSCCKGSRCRGWGESSEDQSLTRIVCAGCGCARCKGPLPDCCHRCAGCT 245

Db 186 TNRSRACHCPCSCCKGSRCRGWGESSEDQSLTRIVCAGCGCARCKGPLPDCCHRCAGCT 245

OY 246 GPKHSCLACLHFHNHSICELHCPALVNTDFFESCPNPEGRTFGASCYVACPYNKL 305

Db 246 GPKHSCLACLHFHNHSICELHCPALVNTDFFESCPNPEGRTFGASCYVACPYNKL 305

OY 306 TDVGSCTLVCPLNQEVTAEDGTORCEKSKPCARGXHSXXPRPAAVPVPPXRQXPXPAHP 365

Db 306 TDVGSCTLVCPLNQEVTAEDGTORCEKSKPCARGXHSXXPRPAAVPVPPXRQXPXPAHP 365

OY 366 VLSFLRPSWDXYSAFYSLPLAPDPDPTSVXISPVSGRQXDPDAHVAV 412

Db 366 VLSFLRPSWDXYSAFYSLPLAPDPDPTSVXISPVSGRQXDPDAHVAV 412

RESULT 5  
 AAE09211 ID AAE09211 standard; Protein; 419 AA.  
 AAE09211; XX  
 15-NOV-2001 (first entry)

XX DE Human p68HER-2 generic protein variant 9.

XX KW HER-2; herstatin; antagonist; extracellular domain; ECD; Herceptin;  
 solid tumour; cancer; polymorphism; cytostatic; gene therapy;

XX KW p68HER-2; ECIIIA; variant.

XX OS Homo sapiens.

XX Key Location/Qualifiers

FT Region 1..340

FT /note= "Identical to N-terminal region of p185HER-2"

FT Domain 341..419

FT /label= ECIIIA variant

FT /note= "Extracellular domain IIIA variant"



PF 16-FEB-2001; 2001WO-US05327.  
 XX  
 PR 16-FEB-2000; 2000US-0506079.  
 XX  
 PA (UYOR-) UNIV OREGON HEALTH SCI.  
 XX  
 PT Clinton G, Henner WD, Evans A;  
 XX  
 DR WPI; 2001-529934/58.

PT New polypeptide, which binds to the extracellular domain of HER-2 for the treatment of hard tumors -  
 XX  
 PS Example 11; Page -; 61pp; English.

XX The invention relates to novel HER-2 (herstatin-2) antagonist particularly a polypeptide that binds to the extracellular domain (ECD) of HER-2 at a site that is different from the binding site of humanised antibody, Herceptin, at an affinity of at least 10<sup>8</sup>. The present invention is based upon the initial discovery of an alternative HER-2 mRNA transcript with 274 bp insert of intron 8. The translation product of the alternative transcript is a truncated HER-2 protein designated p68HER-2 which lacks the transmembrane and intracellular domains of p185HER-2 but contains ECD I, II of the p185HER-2 and the novel ECDIIIA. The ECDIIIA-containing polypeptides bind tightly to, and thus antagonise the HER-2 receptor. The peptides, which bind to an HER-2 ECD, and the nucleic acids encoding these are useful to treat, diagnose and identify solid tumours. The present sequence is human p68HER-2 generic protein Note: The present sequence is not shown in the specification but is derived from p68HER-2 generic sequence (SEQ ID NO:2) shown in the sequence listing (AAE09181).

CC Sequence 419 AA;

XX

Query Match 62.3%; Score 261; DB 22; Length 419;  
 Best Local Similarity 100.0%; Pred. No. 1.6e-237; Indels 0; Gaps 0; Matches 268; Conservative 0; Mismatches 0; Gaps 0;

QY |||||TTPVFTGASPGGLEKLOQLSLPFLCKGGVLIQRNPOLCYQDTIILWKDFHKNNQALTLID 185  
 Db 126 TTPVFTGASPGGLEKLOQLSLPFLCKGGVLIQRNPOLCYQDTIILWKDFHKNNQALTLID 185  
 QY |||||TNRSRAHCPCSPCKGSRCWGSSESDCQLSLTRTCAGGCACRKGPLPTDCCHEQAGCT 245  
 Db 186 TNRSRAHCPCSPCKGSRCWGSSESDCQLSLTRTCAGGCACRKGPLPTDCCHEQAGCT 245  
 QY 246 GKHSDCIACLHRNHSCICELHCPALTYNIDTFESCPNPCCRHYTGASCYTAACPINKL 305  
 Db 246 GPKHSDCIACLHRNHSGICELHCPALTYNIDTFESCPNPCCRHYTGASCYTAACPINKL 305  
 QY 306 TDVGSCILVCPLHNQETAEQDGTCRCKSKPCARGXHSXPRAAVPVPXQXQPXPAHP 365  
 Db 306 TDVGSCILVCPLHNQETAEQDGTCRCKSKPCARGXHSXPRAAVPVPXQXQPXPAHP 365  
 QY 366 VLSFLRPSWDXVSAYFSIPLAPLDPTSV 393  
 Db 366 VLSFLRPSWDXVSAYFSIPLAPLDPTSV 393

RESULT 7

AAE09209 AAE09209 standard; Protein; 419 AA.

AC AAE09209;  
 XX  
 XX 15-NOV-2001 (first entry)

DE Human p68HER-2 generic protein variant 7.

KW HER-2; herstatin; antagonist; extracellular domain; ECD; Herceptin; solid tumour; cancer; polymorphism; cytostatic; gene therapy; p68HER-2; ECDIIIA; variant.

XX

XX Homo sapiens.

OS location/Qualifiers  
 XX 1..340 /note= "identical to N-terminal region of p185HER-2"  
 FH /label= ECDIIIA.variant  
 FT Domain 341..419 /label= ECDIIIA.variant  
 FT /note= "Extracellular domain IIIA variant"  
 FT Misc-difference 124 /note= "Represented as Agn in the parent sequence shown in the specification"  
 FT Misc-difference 125 /note= "Represented as Agn in the parent sequence shown in the specification"  
 FT Misc-difference 342 /label= Unknown  
 FT Misc-difference 345 /label= Unknown  
 FT Misc-difference 346 /label= Unknown  
 FT Misc-difference 356 /label= Unknown  
 FT Misc-difference 358 /label= Unknown  
 FT Misc-difference 361 /label= Unknown  
 FT Misc-difference 376 /label= Unknown  
 FT Misc-difference 394 /note= "p68HER-2 generic sequence (AAE09181) xaa substituted with ile"  
 FT Misc-difference 404 /label= Unknown  
 FT Misc-difference 413 /label= Unknown  
 FT Misc-difference 413 /label= Unknown  
 XX PN WO200161356-A1.  
 XX PD 23-AUG-2001.  
 XX PR 16-FEB-2001; 2001WO-US05327.  
 XX PA (UYOR-) UNIV OREGON HEALTH SCI.  
 PT Clinton G, Henner WD, Evans A;  
 XX DR WPI; 2001-529934/58.

XX The invention relates to novel HER-2 (herstatin-2) antagonist particularly a polypeptide that binds to the extracellular domain (ECD) of HER-2 at a site that is different from the binding site of humanised antibody, Herceptin, at an affinity of at least 10<sup>8</sup>. The present invention is based upon the initial discovery of an alternative HER-2 mRNA transcript with 274 bp insert of intron 8. The translation product of the alternative transcript is a truncated HER-2 protein designated p68HER-2 which lacks the transmembrane and intracellular domains of p185HER-2 but contains ECD I, II of the p185HER-2 and the novel ECDIIIA. The ECDIIIA-containing polypeptides bind tightly to, and thus antagonise the HER-2 receptor. The peptides, which bind to an HER-2 ECD, and the nucleic acids encoding these are useful to treat, diagnose and identify solid tumours. The present sequence is human p68HER-2 generic protein Note: The present sequence is not shown in the specification but is derived from p68HER-2 generic sequence (SEQ ID NO:2) shown in the sequence listing (AAE09181).

SQ	Sequence	419 AA:
FT	Query Match	58.2%; Score 244; DB 22; Length 419;
FT	Best Local Similarity	100.0%; Pred. No. 1.7e-221;
FT	Matches	250; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	126	TPPVIGASPGIRELQLRSLTECLKGVLQIQRNPOLCYQDTILWDIFHKNNQALFLID 185
FT	TTPVIGASPGIRELQLRSLTECLKGVLQIQRNPOLCYQDTILWDIFHKNNQALFLID	185
FT	126	TPPVIGASPGIRELQLRSLTECLKGVLQIQRNPOLCYQDTILWDIFHKNNQALFLID 185
FT	23-AUG-2011.	XX
FT	Misc-difference 413	XX
FT	/label= Unknown	XX
QY	186	TNRSRACHPCSPCCKGSRCWGESESQDQLSLTRVCAGGACRGKGLPPTDCCHCQECAAGCT 245
FT	TNRSRACHPCSPCCKGSRCWGESESQDQLSLTRVCAGGACRGKGLPPTDCCHCQECAAGCT	245
FT	186	TNRSRACHPCSPCCKGSRCWGESESQDQLSLTRVCAGGACRGKGLPPTDCCHCQECAAGCT 245
FT	186	TNRSRACHPCSPCCKGSRCWGESESQDQLSLTRVCAGGACRGKGLPPTDCCHCQECAAGCT 245
FT	GPKHSDCLACLHFNSHGICELHCPLALVNTDTESCPNPPEGRTFGASCWVACPYNL 305	XX
FT	GPKHSDCLACLHFNSHGICELHCPLALVNTDTESCPNPPEGRTFGASCWVACPYNL 305	XX
FT	246	GPKHSDCLACLHFNSHGICELHCPLALVNTDTESCPNPPEGRTFGASCWVACPYNL 305
FT	246	GPKHSDCLACLHFNSHGICELHCPLALVNTDTESCPNPPEGRTFGASCWVACPYNL 305
QY	306	TDVGSCTLVCPHLNQEVTAEDGTCRCKSKCARGHSXXPRPAAVPVPRXQXPXAH 365
FT	TDVGSCTLVCPHLNQEVTAEDGTCRCKSKCARGHSXXPRPAAVPVPRXQXPXAH 365	XX
FT	306	TDVGSCTLVCPHLNQEVTAEDGTCRCKSKCARGHSXXPRPAAVPVPRXQXPXAH 365
FT	306	TDVGSCTLVCPHLNQEVTAEDGTCRCKSKCARGHSXXPRPAAVPVPRXQXPXAH 365
QY	366	VLSFLRPSWD 375
FT	VLSFLRPSWD 375	XX
DB	366	VLSFLRPSWD 375
FT	RESULT 8	XX
FT	AAE09202	CC
FT	AAE09202 standard; Protein: 419 AA.	CC
FT	AAE09202;	CC
FT	AC	CC
FT	XX	CC
FT	XX	CC
FT	15-NOV-2001 (first entry)	CC
FT	Human p68HER-2 generic protein variant (Arg371le).	CC
FT	XX	CC
FT	XX	CC
FT	KW HER-2; herstatin; antagonist; extracellular domain; ECD; Herceptin; solid tumour; cancer; polymorphism; cytostatic; gene therapy; p68HER-2; ECDIIA; variant.	CC
FT	KW	CC
FT	DE	CC
FT	XX	CC
FT	OS Homo sapiens.	CC
FT	XX	CC
FH	Key	Location/qualifiers
FT	Region	1..340
FT	/note= "Identical to N-terminal region of p185HER-2"	
FT	Domain	341..419
FT	/label= ECDIIA.variant	
FT	/note= "Extracellular domain IIIA variant."	
FT	Misc-difference 124	
FT	/note= "Represented as Agn in the parent sequence shown in the specification"	
FT	Misc-difference 125	
FT	/note= "Represented as Agn in the parent sequence shown in the specification"	
FT	Misc-difference 342	
FT	/label= Unknown	
FT	Misc-difference 345	
FT	/label= Unknown	
FT	Misc-difference 346	
FT	/label= Unknown	
FT	Misc-difference 356	
FT	/label= Unknown	
FT	Misc-difference 358	
FT	/label= Unknown	
FT	Misc-difference 361	
FT	/label= Unknown	
FT	Misc-difference 371	
FT	/label= Unknown	
FT	Misc-difference 376	
FT	/note= "p68HER-2 generic sequence (AAE09181) Arg substituted with Ile"	
FT	/label= Unknown	
SQ	Sequence	419 AA:
FT	Query Match	57.0%; Score 239; DB 22; Length 419;
FT	Best Local Similarity	100.0%; Pred. No. 8.8e-217;
FT	Matches	245; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	126	TPPVIGASPGIRELQLRSLTECLKGVLQIQRNPOLCYQDTILWDIFHKNNQALFLID 185
FT	TTPVIGASPGIRELQLRSLTECLKGVLQIQRNPOLCYQDTILWDIFHKNNQALFLID	185
FT	126	TTPVIGASPGIRELQLRSLTECLKGVLQIQRNPOLCYQDTILWDIFHKNNQALFLID 185
FT	126	TTPVIGASPGIRELQLRSLTECLKGVLQIQRNPOLCYQDTILWDIFHKNNQALFLID 185
FT	186	TNRSRACHPCSPCCKGSRCWGESESQDQLSLTRVCAGGACRGKGLPPTDCCHCQECAAGCT 245
FT	TNRSRACHPCSPCCKGSRCWGESESQDQLSLTRVCAGGACRGKGLPPTDCCHCQECAAGCT 245	XX
FT	186	TNRSRACHPCSPCCKGSRCWGESESQDQLSLTRVCAGGACRGKGLPPTDCCHCQECAAGCT 245
FT	186	TNRSRACHPCSPCCKGSRCWGESESQDQLSLTRVCAGGACRGKGLPPTDCCHCQECAAGCT 245
FT	GPKHSDCLACLHFNSHGICELHCPLALVNTDTESCPNPPEGRTFGASCWVACPYNL 305	XX
FT	GPKHSDCLACLHFNSHGICELHCPLALVNTDTESCPNPPEGRTFGASCWVACPYNL 305	XX
FT	246	GPKHSDCLACLHFNSHGICELHCPLALVNTDTESCPNPPEGRTFGASCWVACPYNL 305
FT	246	GPKHSDCLACLHFNSHGICELHCPLALVNTDTESCPNPPEGRTFGASCWVACPYNL 305
FT	246	GPKHSDCLACLHFNSHGICELHCPLALVNTDTESCPNPPEGRTFGASCWVACPYNL 305
FT	306	TDVGSCTLVCPHLNQEVTAEDGTCRCKSKCARGHSXXPRPAAVPVPRXQXPXAH 365
FT	TDVGSCTLVCPHLNQEVTAEDGTCRCKSKCARGHSXXPRPAAVPVPRXQXPXAH 365	XX
FT	306	TDVGSCTLVCPHLNQEVTAEDGTCRCKSKCARGHSXXPRPAAVPVPRXQXPXAH 365
FT	306	TDVGSCTLVCPHLNQEVTAEDGTCRCKSKCARGHSXXPRPAAVPVPRXQXPXAH 365
FT	366	VLSFL 370
FT	366	VLSFL 370
FT	RESULT 9	XX
AAE09207		XX

ID AAE09207 standard; Protein; 419 AA.  
 XX  
 AC p68HER-2;  
 XX  
 15-NOV-2001 (first entry)  
 DT  
 DE Human p68HER-2 generic protein variant 5.  
 KW HER-2; herstatin; antagonist; extracellular domain; ECD; Herceptin;  
 KW solid tumour; cancer; polymorphism; cytostatic; gene therapy;  
 KW p68HER-2; ECDIIIA; variant.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Region 1..340  
 /note= "Identical to N-terminal region of p185HER-2"  
 FT Domain 341..419  
 /label= ECDIIIA.variant  
 /note= "Extracellular domain IIIA variant."  
 FT Misc-difference 124  
 /note= "Represented as Agn in the parent sequence shown in the specification"  
 FT Misc-difference 125  
 /note= "Represented as Agn in the parent sequence shown in the specification"  
 FT Misc-difference 342  
 /label= Unknown  
 FT Misc-difference 345  
 /label= Unknown  
 FT Misc-difference 346  
 /label= Unknown  
 FT Misc-difference 356  
 /label= Unknown  
 FT Misc-difference 358  
 /note= "p68HER-2 generic sequence (AAE09181) Xaa substituted with Leu"  
 FT Misc-difference 361  
 /label= Unknown  
 FT Misc-difference 376  
 /label= Unknown  
 FT Misc-difference 394  
 /label= Unknown  
 FT Misc-difference 404  
 /label= Unknown  
 FT Misc-difference 413  
 /label= Unknown  
 PN WO2001151356-A1.  
 XX  
 PD 23-AUG-2001.  
 XX  
 PF 16-FEB-2001; 2001WO-US05327.  
 XX  
 PR 16-FEB-2000; 2000US-0506079.  
 XX  
 PA (UYOR-) UNTV OREGON HEALTH SCI.  
 PI Clinton G., Henner WD., Evans A.;  
 XX  
 DR WPI; 2001-529934/58.  
 XX  
 PT New polypeptide, which binds to the extracellular domain of HER-2 for the treatment of hard tumors -  
 XX  
 PS Example 11; Page - ; 61pp; English.

The invention relates to novel HER-2 (herstatin-2) antagonist particularly a polypeptide that binds to the extracellular domain (ECD) of HER-2 at a site that is different from the binding site of humanised antibody, Herceptin, at an affinity of at least 10<sup>8</sup>. The present invention is based upon the initial discovery of an alternative HER-2 mRNA transcript with 274 bp insert of intron 8. The translation product

CC of the alternative transcript is a truncated HER-2 protein designated p68HER-2 which lacks the transmembrane and intracellular domains of p185HER-2 but contains ECD I, II of the p185HER-2 and the novel ECDIIIA. The ECDIIIA-containing polypeptides bind tightly to, and thus antagonise the HER-2 receptor. The peptides, which bind to an HER-2 ECD, and the nucleic acids encoding these are useful to treat, diagnose and identify solid tumours. The present sequence is human p68HER-2 generic protein containing ECDIIIA variant sequence.  
 Note: The present sequence is not shown in the specification but is derived from p68HER-2 generic sequence (SBO ID NO:2) shown in the sequence listing (AAE09181).

XX Sequence 419 AA:

Query	Match	Score	DB	Length	
Best Local Similarity	100.0%	Pred. No.	2.1e-206	419	
Matches	232	Conservative	0	Mismatches	0
Indels	0	Gaps	0		

Db QY 126 TTVVTGASPGGLELQLRSLTICLKGTYLQRNPOLCYQDTTLWKDIFHKNNQALTLLID 185  
 Db 126 TTPVTGASPGGLELQLRSLTICLKGTYLQRNPOLCYQDTTLWKDIFHKNNQALTLLID 185  
 Db QY 186 TARSRACHPCSPCCKGSRKWCSESSEDQSLTRNPGVAGCARGACKGPIPTCQHQCAGCT 245  
 Db 186 TNKSRAHCPCSPCCKGSRKWCSESSEDQSLTRNPGVAGCARGACKGPIPTCQHQCAGCT 245  
 Qy 246 GRKHSDCIACLHFNHNSGICELHCPALVYTIDFESCPNPPEGRTYFGASCYTCAPYNKLS 305  
 Db 246 GRKHSDCIACLHFNHSGICELHCPALVYTIDFESCPNPPEGRTYFGASCYTCAPYNKLS 305  
 Db QY 306 TDVGSCILVCPLHNQEVTEDGTORCEKSKFCARGXHSXXPRPAAVPVPX 357  
 Db 306 TDVGSCILVCPLHNQEVTEDGTORCEKSKFCARGXHSXXPRPAAVPVPX 357

! RESULT 10  
 AR09183  
 ID AAE09183 standard; Protein; 419 AA.  
 XX  
 AC AAE09183;  
 XX  
 DT 15-NOV-2001 (first entry)  
 XX  
 DE Human p68HER-2 generic sequence #2.  
 KW HER-2; herstatin; antagonist; extracellular domain; ECD; Herceptin;  
 KW solid tumour; cancer; polymorphism; cytostatic; gene therapy;  
 KW p68HER-2; ECDIIIA.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Region 1..340  
 /note= "Identical to N-terminal region of p185HER-2"  
 FT Domain 341..419  
 /label= ECDIIIA  
 /note= "Extracellular domain IIIA"  
 FT Misc-difference 124  
 /note= "Represented as Agn in the sequence shown in the specification"  
 FT Misc-difference 125  
 /note= "Represented as Agn in the sequence shown in the specification"  
 FT Misc-difference 342  
 /label= Unknown  
 FT Misc-difference 345  
 /note= "Encoded by WCC"  
 FT Misc-difference 346  
 /note= "Encoded by CYG"  
 FT Misc-difference 347  
 /label= Unknown  
 FT Misc-difference 348  
 /note= "Encoded by CYC"

FT /label= Unknown  
 FT /note= "Encoded by CWG"  
 FT 357  
 FT /label= Unknown  
 FT /note= "Encoded by YGC"  
 FT 358  
 FT /label= Unknown  
 FT /note= "Encoded by ATR"  
 FT 361  
 FT /label= Unknown  
 FT /note= "Encoded by GNC"  
 FT 371  
 FT /label= Unknown  
 FT /note= "Encoded by AKA"  
 FT 376  
 FT /label= Unknown  
 FT /note= "Encoded by MTA"  
 FT 389  
 FT /label= Unknown  
 FT /note= "Encoded by AGC"  
 FT 394  
 FT /label= Unknown  
 FT /note= "Encoded by CST"  
 FT 404  
 FT /label= Unknown  
 FT /note= "Encoded by CYG"  
 FT 413  
 FT /label= Unknown  
 FT /note= "Encoded by VAC"  
 XX  
 PN WO200161356-A1.  
 XX  
 PD 23-AUG-2001.  
 XX  
 PF 16-FEB-2001; 2001WO-US05327.  
 XX  
 PR 16-FEB-2000; 2000US-0506079.  
 XX  
 PA (UOR-) UNIV OREGON HEALTH SCI.  
 XX  
 PI Clinton G, Henner WD, Evans A;  
 XX  
 DR WPI; 2001-529934/58.  
 XX  
 DR N-PSDB; AAD15852.  
 XX  
 PT New polypeptide, which binds to the extracellular domain of HER-2 for  
 PT the treatment of hard tumors  
 PT  
 PS Claim 8; Page 57-58; 61pp; English.  
 XX  
 CC The invention relates to novel HER-2 (herstatin-2) antagonist  
 CC particularly a polypeptide that binds to the extracellular domain (ECD)  
 CC of HER-2 at a site that is different from the binding site of humanised  
 CC antibody, Herceptin, at an affinity of at least 10<sup>8</sup>. The present  
 CC invention is based upon the initial discovery of an alternative HER-2  
 CC mRNA transcript with 274 bp insert of intron 8. The translation product  
 CC of the alternative transcript is a truncated HER-2 protein designated  
 CC p68HER-2 which lacks the transmembrane and intracellular domains of  
 CC p18HER-2 but contains ECD I, II of the p18HER-2 and the novel ECDIIIA.  
 CC The ECDIIIA-containing polypeptides bind tightly to, and thus antagonise  
 CC the HER-2 receptor. The peptides, which bind to an HER-2 ECD, and the  
 CC nucleic acids encoding these are useful to treat, diagnose and identify  
 CC solid tumours. The present sequence is human p68HER-2 protein  
 CC containing ECDIIIA generic sequence.  
 XX  
 SQ sequence 419 AA;

Query Match 54.2%; Score 227; DB 22; Length 419;  
 Best Local Similarity 100.0%; Pred. No. 1.8e-205;  
 Matches 230; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 126 TPTVIGASPGGLELQLRSLTECLKGGLIQRNPOLCYOPTILWDIFHKNNQLALTLID 185  
 DB 126 TPTVIGASPGGLELQLRSLTECLKGGLIQRNPOLCYOPTILWDIFHKNNQLALTLID 185

QY 186 TNRSRACHCSPCCCKGSRQWGESESDCQSLRTVQAGGCARKGPLPTDCCHEOCAGGT 245  
 DB 186 TNRSRACHCSPCCCKGSRQWGESESDCQSLRTVQAGGCARKGPLPTDCCHEOCAGGT 245  
 QY 246 GPKHSDCLACLHFNHSGICELHCPLAVTYNDTEFSCPPPEGRYTFGACSVTAPYNKLS 305  
 DB 246 GPKHSDCLACLHFNHSGICELHCPLAVTYNDTEFSCPPPEGRYTFGACSVTAPYNKLS 305  
 QY 306 TDVGCTLWPLHNGEVTAEDGTQCEKCSKPCARGXHSXKPRPAVVP 355  
 DB 306 TDVGCTLWPLHNGEVTAEDGTQCEKCSKPCARGXHSXKPRPAVVP 355

RESULT 11  
 AAE09200  
 ID AAE09200 standard; Protein; 419 AA.  
 AC AAE09200;  
 XX  
 DT 15-NOV-2001 (first entry)  
 DE Human p68HER-2 generic protein variant (Arg9357Cys).  
 KW HER-2; herstatin; antagonist; extracellular domain; ECD; Herceptin;  
 KW solid tumour; cancer; polymorphism; cytostatic; gene therapy;  
 KW p68HER-2; ECDIIIA; variant.  
 XX  
 OS Homo sapiens.  
 XX  
 FR Key Location/Qualifiers  
 FT Region 1..340  
 FT /note= "Identical to N-terminal region of p185HER-2"  
 FT Domain 341..419  
 FT /label= ECDIIIA\_variant  
 FT /note= "Extracellular domain IIIa variant"  
 FT Misc-difference 124  
 FT /note= "Represented as Agn in the parent sequence shown  
 FT in the specification"  
 FT Misc-difference 125  
 FT /note= "Represented as Agn in the parent sequence shown  
 FT in the specification"  
 FT Misc-difference 342  
 FT /label= Unknown  
 FT Misc-difference 345  
 FT /label= Unknown  
 FT Misc-difference 346  
 FT /label= Unknown  
 FT Misc-difference 356  
 FT /label= Unknown  
 FT Misc-difference 357  
 FT /note= "p68HER-2 generic sequence (AAP09181) Arg  
 FT substituted with Cys"  
 FT Misc-difference 358  
 FT /label= Unknown  
 FT Misc-difference 361  
 FT /note= "Label= Unknown  
 FT Misc-difference 376  
 FT /label= Unknown  
 FT Misc-difference 394  
 FT /label= Unknown  
 FT Misc-difference 404  
 FT /label= Unknown  
 FT Misc-difference 413  
 FT /label= Unknown  
 XX  
 PN WO200161356-A1.

XX  
 PD 23-AUG-2001.  
 XX  
 DT 16-FEB-2001; 2001WO-US05327.  
 XX  
 PR 16-FEB-2000; 2000US-0506079.

XX (UYOR-) UNIV OREGON HEALTH SCI.  
 XX PA /label= ECDIIIA variant  
 XX FT /note= "Extracellular domain IIIA variant"  
 PI Clinton G., Henner WD, Evans A;  
 XX FT Misc-difference 124 /note= "Represented as Agn in the parent sequence shown  
 DR XX FT Misc-difference 125 /note= "Represented as Agn in the parent sequence shown  
 XX PT the treatment of hard tumors - in the specification"  
 XX PS Example 12; Page -; 61PP; English.  
 XX SQ Sequence 419 AA:  
 Query Match 54.2%; Score 227; DB 22; Length 419;  
 Best Local Similarity 100.0%; Pred. No. 1. 8e-205; Mismatches 0; Indels 0; Gaps 0;  
 Matches 230; Conservative 0; MisMatches 0; Indels 0; Gaps 0;  
 CC 126 TTPVTGASPGGLELQLRSLSFLCCKLGVLQINPQLCYQDFTLWKFHKNNQALATLID 185  
 Db 126 TTPVTGASPGGLELQLRSLSFLCCKLGVLQINPQLCYQDFTLWKFHKNNQALATLID 185  
 CC 186 TNRSRACHPCSPCKCSKRCWSSESDQSLTRVCAAGGCGARCKGPLTDCHEQAGCT 245  
 Db 186 TNSRACHPCSPCKGSRCWGESESDQSLTRVCAAGGCGARCKGPLTDCHEQAGCT 245  
 CC 246 GPKHSDCLACIHFNHSCICELHCPCALVYNTDFESCPNPNSGRYTGASCYTACPNKL 305  
 Db 246 GPKHSDCLACIHFNHSCICELHCPCALVYNTDFESCPNPNSGRYTGASCYTACPNKL 305  
 CC 306 TDVGSCILVCPHLNQEVTAEDGTQERKCSAPRGHXSXPRPAVPVP 355  
 Db 306 TDVGSCILVCPHLNQEVTAEDGTQERKCSAPRGHXSXPRPAVPVP 355  
 RESULT 12  
 AAE09206 ID AAE09206 standard; Protein: 419 AA.  
 XX AC AAE09206;  
 XX DT 15-NOV-2001 (first entry)  
 XX DE Human p68HER-2 generic protein variant 4.  
 XX HER-2; herstatin; antagonist; extracellular domain; ECD; Herceptin;  
 KW solid tumour; cancer; polymorphism; cytostatic; gene therapy;  
 KW p68HER-2; ECDIIIA; variant.  
 XX OS Homo sapiens.  
 XX SQ Sequence 419 AA:  
 Query Match 54.2%; Score 227; DB 22; Length 419;  
 Best Local Similarity 100.0%; Pred. No. 1. 8e-205; Mismatches 0; Indels 0; Gaps 0;  
 Matches 230; Conservative 0; MisMatches 0; Indels 0; Gaps 0;  
 CC 341.419 FT Key Region 1..340 /note= "Identical to N-terminal region of p185HER-2"  
 FT Domain

QY 126 TTPVPGASPGGLREQLRSITTECLKGVLIRNPOLCYQDTILWQDIFHNNQALTLID 185  
 ||||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 126 TTPVPGASPGGLREQLRSITTECLKGVLIRNPOLCYQDTILWQDIFHNNQALTLID 185  
 XX  
 QY 186 TNRSRACHPCSPCCGKGSRCMGESSEDQSLTRTVAAGGCARCKGGLPPTDCHEQCAAGCT 245  
 ||||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 186 TNRSRACHPCSPCCGKGSRCMGESSEDQSLTRTVAAGGCARCKGGLPPTDCHEQCAAGCT 245  
 PA (UYOR-) UNIV OREGON HEALTH SCI.  
 XX  
 QY 246 GPKHSDCLACLHFHNHGICELHCPLAVTYNNTFESCPNPBEGRTFGASCYTACPYNLK 305  
 ||||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 246 GPKHSDCLACLHFHNHGICELHCPLAVTYNNTFESCPNPBEGRTFGASCYTACPYNLK 305  
 PR 16-FEB-2000; 2000US-0506079.  
 DR WPI; 2001-529934/58.  
 XX  
 QY 306 TDVGSCTLVCPLNHQEVTAEDGTQRCEKCSKPCARGXHSXPRPAVVP 355  
 ||||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 306 TDVGSCTLVCPLNHQEVTAEDGTQRCEKCSKPCARGXHSXPRPAVVP 355  
 PT New polypeptide, which binds to the extracellular domain of HER-2 for  
 XX the treatment of hard tumors -  
 PS Example 12; Page -; 61pp; English.  
 XX  
 CC The invention relates to novel HER-2 (herstatin-2) antagonist  
 CC particularly a polypeptide that binds to the extracellular domain (ECD)  
 CC of HER-2 at a site that is different from the binding site of manumised  
 CC antibody, Herceptin, at an affinity of at least 10<sup>-8</sup>. The present  
 CC invention is based upon the initial discovery of an alternative HER-2  
 CC mRNA transcript with 274 bp insert of intron 8. The translation product  
 CC of the alternative transcript is a truncated HER-2 protein designated  
 CC p68HER-2 which lacks the transmembrane and intracellular domains of  
 CC p185HER-2 but contains ECD I, II of the p185HER-2 and the novel ECDIIA.  
 CC The ECDIIA-containing polypeptides bind tightly to, and thus antagonise  
 CC the HER-2 receptor. The peptides, which bind to an HER-2 ECD, and the  
 CC nucleic acids encoding these are useful to treat, diagnose and identify  
 CC solid tumours. The present sequence is a truncated HER-2 protein designated  
 CC containing ECDIIA variant sequence.  
 CC Note: The present sequence is not shown in the specification but is  
 CC derived from p68HER-2 generic sequence (SEQ ID NO:13) shown in the  
 CC sequence listing (AAE09183).  
 XX  
 SQ Sequence 419 AA;  
 Query Match 54.2%; Score 227; DB 22; Length 419;  
 Best Local Similarity 10.0%; Pred. No. 1.8e-205; Matches 230; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 FT /note= "represented as Agn in the sequence shown in  
 FT the specification"  
 FT Misc-difference 125 /note= "represented as Agn in the sequence shown in  
 the specification"  
 FT Misc-difference 342 /label= Unknown  
 FT Misc-difference 345 /label= Unknown  
 FT Misc-difference 346 /label= Unknown  
 FT Misc-difference 356 /label= Unknown  
 FT Misc-difference 357 /note= "p68HER-2 generic sequence (AAE09183) xaa  
 substituted with Cys"  
 FT Misc-difference 358 /label= Unknown  
 FT Misc-difference 361 /label= Unknown  
 FT Misc-difference 371 /label= Unknown  
 FT Misc-difference 376 /label= Unknown  
 FT Misc-difference 394 /label= Unknown  
 FT Misc-difference 404 /label= Unknown  
 FT Misc-difference 413 /label= Unknown  
 PN WO200161356-A1.  
 OS Homo sapiens.

---

PD 23-AUG-2001.  
 XX  
 PF 16-FEB-2001; 2001WO-US05327.  
 XX  
 AC AAE09215;  
 XX  
 DT 15-NOV-2001 (first entry)  
 XX  
 XX Human p68HER-2 generic protein variant 14.  
 XX HER-2; heparin; antagonist; extracellular domain; ECD; Herceptin;  
 XX solid tumour; cancer; polymorphism; cytostatic; gene therapy;  
 XX p68HER-2; ECDIIA; variant.  
 XX  
 OS Homo sapiens.

XX	FH	Location/Qualifiers	CC	derived from P68HER-2 generic sequence (SBO ID NO:13) shown in the sequence listing (AAE09183).
FT	Region	1..340	CC	
FT	/note= "Identical to N-terminal region of p185HER-2"		SQ	
FT	Domain	341..419		
FT	/label= ECDIIIA variant			
FT	/note= "Extracellular domain IIIa."			
FT	/note= "Represented as Arg in the sequence shown in the specification"			
FT	Misc-difference 342			
FT	/note= "Label= Unknown			
FT	Misc-difference 345			
FT	/note= "Represented as Arg in the sequence shown in the specification"			
FT	Misc-difference 346			
FT	/note= "Label= Unknown			
FT	Misc-difference 356			
FT	/note= "Label= Unknown			
FT	Misc-difference 357			
FT	/note= "Label= Unknown			
FT	Misc-difference 358			
FT	/note= "Label= Unknown			
FT	Misc-difference 361			
FT	/note= "Label= Unknown			
FT	Misc-difference 371			
FT	/note= "Label= Unknown			
FT	Misc-difference 376			
FT	/note= "Label= Unknown			
FT	Misc-difference 394			
FT	/note= "Label= Unknown			
FT	Misc-difference 404			
FT	/note= "Label= Unknown			
FT	Misc-difference 413			
FT	/note= "Label= Unknown			
FT	W0200161356-A1.			
XX	PD	23-AUG-2001.		
XX	PF	16-FEB-2001; 2001WO-US05327.		
XX	PR	16-FEB-2000; 2000US-0506079.		
XX	PA	(UYOR-) UNIV OREGON HEALTH SCI.		
XX	PT	Clinton, G., Henner WD, Evans A;		
XX	DR	WPI; 2001-529934/58.		
XX	PS	Example 12; Page - ; 61pp; English.		
CC	The invention relates to novel HER-2 (herstatin-2) antagonist polypeptides, particularly a polypeptide that binds to the extracellular domain (ECD) of HER-2 at a site that is different from the binding site of humanised antibody, Herceptin, at an affinity of at least 10 <sup>-8</sup> . The present invention is based upon the initial discovery of an alternative HER-2 mRNA transcript with 274 bp insert of intron 8. The translation product of the alternative transcript is a truncated HER-2 protein designated p68HER-2 which lacks the transmembrane and intracellular domains of p185HER-2 but contains ECD I, II of the p185HER-2 and the novel ECDIIIA. The ECDIIIA-containing polypeptides bind tightly to, and thus antagonise the HER-2 receptor. The peptides, which bind to an HER-2 ECD and the nucleic acids encoding these are useful to treat, diagnose and identify solid tumours. The present sequence is human p68HER-2 generic protein containing ECDIIIA variant sequence.			
CC	Note: The present sequence is not shown in the specification but is			
XX	Key	Location/Qualifiers	Query Match	Best local Similarity 100.0%; Score 227; DB 22; Length 419;
FT	Region	1..340	Matches 230; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
FT	/note= "Identical to N-terminal region of p185HER-2"			
FT	Domain	341..419	Sequence 419 AA.	
FT	/label= ECDIIIA variant			
FT	/note= "Extracellular domain IIIa."			
FT	/note= "Represented as Arg in the sequence shown in the specification"			
FT	Misc-difference 342			
FT	/note= "Label= Unknown			
FT	Misc-difference 345			
FT	/note= "Represented as Arg in the sequence shown in the specification"			
FT	Misc-difference 346			
FT	/note= "Label= Unknown			
FT	Misc-difference 356			
FT	/note= "Label= Unknown			
FT	Misc-difference 357			
FT	/note= "Label= Unknown			
FT	Misc-difference 358			
FT	/note= "Label= Unknown			
FT	Misc-difference 361			
FT	/note= "Label= Unknown			
FT	Misc-difference 371			
FT	/note= "Label= Unknown			
FT	Misc-difference 376			
FT	/note= "Label= Unknown			
FT	Misc-difference 394			

PT Misc-difference /label= Unknown 404  
 PT Misc-difference /label= Unknown 413  
 PT note "p68HER-2 generic sequence (AAB09183) xaa  
 PT substituted w/Asn"  
 XX WO200161356-A1.  
 XX PD 23-AUG-2001.  
 XX PF 16-FEB-2001; 2001WO-US05327.  
 XX PR 16-FEB-2000; 2000US-0506079.  
 XX PA (UYOR-) UNIV OREGON HEALTH SCI.  
 XX Clinton G, Henner WD, Evans A;  
 DR WPI; 2001-529934/58.  
 XX PT New polypeptide, which binds to the extracellular domain of HER-2 for  
 PT the treatment of hard tumors -  
 XX PS Example 12: Page -; 61pp; English.

The invention relates to novel HER-2 (herstatin-2) antagonist  
 CC particularly a polypeptide that binds to the extracellular domain (ECD)  
 CC of HER-2 at a site that is different from the binding site of humanised  
 CC antibody, Herceptin, at an affinity of at least 10<sup>-8</sup>. The present  
 CC invention is based upon the initial discovery of an alternative HER-2  
 CC mRNA transcript with 274 bp insert of intron 8. The translation product  
 CC of the alternative transcript is a truncated HER-2 protein designated  
 CC p68HER-2 which lacks the transmembrane and intracellular domains of  
 CC p188HER-2 but contains ECD I, II of the p188HER-2 and the novel ECDIIIA.  
 CC The ECDIIIA-containing polypeptides bind tightly to, and thus antagonise  
 CC the HER-2 receptor. The peptides, which bind to an HER-2 ECD, and the  
 CC nucleic acids encoding these are useful to treat, diagnose and identify  
 CC solid tumours. The present sequence is human p68HER-2 generic protein  
 CC containing ECDIIIA variant sequence.  
 Note: The present sequence is not shown in the specification but is  
 CC derived from p68HER-2 generic sequence (SEQ ID NO:13) shown in the  
 CC sequence listing (RAE09183).

XX Sequence 419 AA;

Query Match 54.2%; Score 227; DB 22; Length 419;  
 Best Local Similarity 100.0%; Pred. No. 1.8e-205;  
 Matches 230; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 126 TTPVYGASPOGLREQLRSITCECLKGGVLIQRNPOLCYQDTIILWMDIFHKNQNQALTLID 185
Db 126 TTPVYGASPOGLREQLRSITCECLKGGVLIQRNPOLCYQDTIILWMDIFHKNQNQALTLID 185
QY 186 TNSRSTRACPHCSPCCKGSRGEGESSDCQSLTRTCAGGCACKGGLPPTDCCHEOCAGCT 245
Db 186 TNSRSTRACPHCSPCCKGSRGEGESSDCQSLTRTCAGGCACKGGLPPTDCCHEOCAGCT 245
QY 246 GPKHSDCCLACLHFNHSGICELHCPALVTNTDFESCPNPEGRRTFGASCVTAZYPKLS 305
Db 246 GPKHSDCCLACLHFNHSGICELHCPALVTNTDFESCPNPEGRRTFGASCVTAZYPKLS 305
QY 306 TDVGSCCTLVPLHNOEVTAEDGTQRCCKSKPCARGXHSXPRPAVPV 355
Db 306 TDVGSCCTLVPLHNOEVTAEDGTQRCCKSKPCARGXHSXPRPAVPV 355

```

Search completed: January 14, 2003, 17:12:55  
 Job time : 54.8474 secs

**THIS PAGE BLANK (USPTO)**

Copyright (c) 1993 - 2003 Compugen Ltd.

Gencore version 5.1.3

OM protein - protein search, using sw model  
Run on: January 14, 2003, 17:10:43 ; Search time 8.7249 seconds  
Sequence: (without alignments)  
Scoring table: 1865.663 Million cell updates/sec

## Title:

US-09-506-079E-1  
Perfect score: 79  
Sequence: GAXHSXXPRPAAVPVVPRXQP.....VGRGXDPDAHVAVKLSRYEG 79

Scored: 671580 seqs, 206047115 residues

Word size : 0

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : SPTREMBL\_21:  
1: sp\_archaea:  
2: sp\_bacteria:  
3: sp\_fungi:  
4: sp\_human:  
5: sp\_invertebrate:  
6: sp\_mammal:  
7: sp\_micr:  
8: sp\_organelle:  
9: sp\_phage:  
10: sp\_plant:  
11: sp\_rabbit:  
12: sp\_virus:  
13: sp\_vertebrate:  
14: sp\_unclassified:  
15: sp\_xvirus:  
16: sp\_bacteriaph:  
17: sp\_archeap:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	17	21.5	419 4 Q9UK79	Q9uk79 homo sapien
2	7	8.9	27 4 O9HD19	O9hd19 homo sapien
3	7	8.9	121 16 Q8YKWB	Q8ykwb anaebena sp
4	7	8.9	178 10 Q9FMC7	Q9fmc7 arabidopsis
5	7	8.9	245 4 O14596	O14596 homo sapien
6	7	8.9	268 2 Q9EWB9	Q9ewb9 streptomyce
7	7	8.9	382 5 Q9U3EL	Q9u3el caenorhabdi
8	7	8.9	389 16 Q8UEF8	Q8uef8 agrobacteri
9	7	8.9	416 16 Q9KZK5	Q9kzk5 streptomyce
10	7	8.9	436 16 O67814	O67814 aquifex aeo
11	7	8.9	469 5 Q8WQ63	Q8wq63 drosophila
12	7	8.9	469 5 Q8WPA9	Q8wp9 drosophila
13	7	8.9	469 5 Q8WP54	Q8wp54 drosophila
14	7	8.9	471 5 Q8WQ65	Q8wq65 drosophila
15	7	8.9	471 5 Q8WQ64	Q8wq64 drosophila
16	7	8.9	471 5 Q8WP56	Q8wp56 drosophila

Database : SPTREMBL\_21:  
1: sp\_archaea:  
2: sp\_bacteria:  
3: sp\_fungi:  
4: sp\_human:  
5: sp\_invertebrate:  
6: sp\_mammal:  
7: sp\_micr:  
8: sp\_organelle:  
9: sp\_phage:  
10: sp\_plant:  
11: sp\_rabbit:  
12: sp\_virus:  
13: sp\_vertebrate:  
14: sp\_unclassified:  
15: sp\_xvirus:  
16: sp\_bacteriaph:  
17: sp\_archeap:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## ALIGNMENTS

RESULT	ID	PRELIMINARY;	PRT;	419 AA.
1	Q9UK79			
	Q9UK79;			
	DT 01-MAY-2000 (TREMBLrel. 13, Created)			
	DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)			
	DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)			
	DE Heratin.			
	GN HER-2.			
	OS Homo sapiens (Human).			
	OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
	NCBI-TaxID=9606;			
	RN [1]			
	RP SEQUENCE FROM N.A.			
	RX MEDLINE=99415951; PubMed=10405918;			
	RA Doherry J.K., Bond C., Jardim A., Adelman J.P., Clinton G.M.;			
	RT "The HER-2/neu receptor tyrosine kinase gene encodes a secreted autoinhibitor."			
	RT Proc. Natl. Acad. Sci. U.S.A. 96:10869-10874(1999).			
	RP [2]			
	RP SEQUENCE FROM N.A.			
	RA Doherry J.K., Clinton G.M., Adelman J.P., Evans A.J., Henner W.D.;			
	RL Submitted (SEP-2000) to the EMBL/GenBank/DDBJ databases.			
	DR AF17761; AA00909; -			
	DR InterPro; IPR000494; EGFR_L-domain.			
	DR InterPro; IPR002174; Furin-like.			
	DR Pfam; PF00757; Furin-like; 1.			
	DR Pfam; PF01030; Reccep_L-domain; 1.			
	DR SMART; SM00261; FU; 1.			
	SQ SEQUENCE 419 AA; 45472 MW; FEC1BE347E2D030C CRC64;			
	Query Match 21.5%; Score 17; DB 4; Length 419;			
	Best Local Similarity 10.0%; Pred. No. 1.9e-09;			
	Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
	OY 37 VSAYSLPLAPLSPSPV 53			
	Db .377 VSAYSLPLAPLSPSPV 393			



DT 01-JAN-1998 (TREMBLrel. 05, Created)  
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)  
 DE Folate binding protein.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 NCBI\_TaxID=9605;  
 OX RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=SALIVARY GLAND;  
 RA Verma R.S., Elwood P.C.;  
 RT "Identification of a homologous cDNA to folate receptor from human  
 salivary gland.", Submitted (APR-1997) to the EMBL/GenBank/DDJB databases.  
 RL EMBL: AF000380; AAB81937; 1;  
 DR IPR004259; Folate\_rec.  
 DR Pfam: PF03024; Folate\_rec; 1.  
 SQ SEQUENCE 245 AA; 26268 MW; BA36B7DB95FC0E820 CRC64;  
 Query Match 8.9%; Score 7; DB 4; Length 245;  
 Best Local Similarity 100.0%; Pred. No. 28;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 46 APLSPRS 52  
 Db 187 APLSPRS 193

RESULT 6  
 Q9EW99 PRELIMINARY; PRT; 268 AA.  
 AC Q9EW99 ID DT 01-MAR-2001 (TREMBLrel. 16, Created)  
 OC Actinomycetales; Streptomyces. DE 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
 RN NCBL\_TAXID=1911; DE ABC-transporter.  
 OS Streptomyces griseus.  
 OC Bacteria; Firmicutes; Actinobacteridae; Actinobacteridae;  
 OC Actinomycetales; Streptomyceae; Streptomyces.  
 RN STRAIN=IMRU 3570; RA Campelo A.B., Gil J.A.;  
 RT "Cloning and characterization of a gene cluster from Streptomyces  
 griseus IMRU 3570 involved in candicidin production.", Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.  
 RL [2] SEQUENCE FROM N.A.  
 RC STRAIN=IMRU 3570;  
 RA Campelo A.B.; Thesis (2000); Department of Microbiologia, Universidad de Leon, .  
 DR EMBL: AJ30302; CAC21191; -  
 DR InterPro: IPR00412; ABCtransporter.  
 DR Pfam: PF01061; ABC2\_membrane; 1.  
 DR PRINTS: PRO0164; ABC2TRANSPORT.  
 DR PROSITE: PS00590; ABC2\_MEMBER; UNKNOWN 1.  
 SQ SEQUENCE 268 AA; 27945 MW; C621E71D2EFTB4 CRC64;  
 Query Match 8.9%; Score 7; DB 2; Length 268;  
 Best Local Similarity 100.0%; Pred. No. 30;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 PRPAAVP 13  
 AC Q9U3E1 ID DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 OC Eukaryota; Metazoa; Nematoidea; Rhabditida; Rhabditoidea;  
 OC Rhabditidae; Peioderinae; Caenorhabditis. DE F58A4\_7B protein.  
 RN NCBL\_TAXID=6239; RX [1]  
 RP SEQUENCE FROM N.A.  
 RA Berks M.; Submitted (APR-1993) to the EMBL/GenBank/DDJB databases.  
 RL EMBL: AF000380; AAB81937; 1;  
 DR MEDLINE=90069613; PubMed=9851916;  
 RA none;  
 RT "Genome sequence of the nematode *C.elegans*: A platform for  
 investigating biology.", Submitted (APR-1993) to the EMBL/GenBank/DDJB databases.  
 RL Science 283:2012-2018(1998).  
 DR EMBL: Z22179; CAA80170; 1;  
 DR HSSP; P36556; IAM9.  
 DR InterPro: IPR001092; HLH basic.  
 DR Pfam: PF00010; HLH; 1.  
 DR SMART; SM0035; HLH; 1.  
 DR PROSITE; PS00038; HELIX\_LOOP\_HELIX; UNKNOWN 1.  
 DR SEQUENCE 382 AA; 40763 MW; 510BEF225B073804 CRC64;  
 Query Match 8.9%; Score 7; DB 5; Length 382;  
 Best Local Similarity 100.0%; Pred. No. 40;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 46 APLSPRS 52  
 Db 48 APLSPRS 54

RESULT 8  
 Q8UEFB PRELIMINARY; PRT; 389 AA.  
 AC Q8UEFB ID DT 01-JUN-2002 (TREMBLrel. 21, Created)  
 OC Agrobacterium tumefaciens (strain C58 / ATCC 33970). DE 01-JUN-2002 (TREMBLrel. 21, Last sequence update)  
 RN KU1800 OR AGRC\_3311. OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).  
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group; OC Rhizobiaceae; Rhizobium.  
 OC NCBI\_TAXID=176299; RX [1]  
 RP SEQUENCE FROM N.A. RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,  
 RA MEDLINE=21608550; PubMed=11743193;  
 RA Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,  
 RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Bovée D. Sr.,  
 RA Chapman P., Clendinning J., Deathridge G., Gillet W., Grant C.,  
 RA Kuttyarin T., Levy R., Li M.-J., McClelland E., Palmeri A.,  
 RA Raymond C., Rouse G., Saenphimachak C., Wu Z., Romero P., Gordon D.,  
 RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Kreepan W., Perry M.,  
 RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,  
 RA Chumley F., Tingey S. V., Tomb J.-F., Gordon M.P., Olson M.V.,  
 RA Nester E.W.; RT "The genome of the natural genetic engineer Agrobacterium tumefaciens  
 C58.", Science 294:2317-2323(2001).  
 RL RN [2]  
 RP SEQUENCE FROM N.A. RX MEDLINE=21608551; PubMed=11743194;  
 RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,  
 RA Qurollo B., Goldman B.S., Cao Y., Askrenzi M., Halling C., Mullin L.,  
 RA Houmied K., Gordon J., Vaudin M., Iartchouk O., Epp A., Liu F.,

RESULT 7  
 Q9U3E1 PRELIMINARY; PRT; 382 AA.  
 ID Q9U3E1

		Matches	7;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
RA	Wollam C., Allinder M., Doughty D., Scott C., Lappas C., Markelz B.,										
RA	Finnegan C., Crowell C., Gursen J., Lomo C., Sear C., Strub G.,										
RA	Cielo C., Slater S.,										
RT	"Genome sequence of the plant pathogen and biotechnology agent										
RT	aerobacterium tumefaciens C58.";										
RL	Science 294:2323-2328 (2001).										
DR	EMBL; AB009135; AAC42198.1; -.										
KW	Hypothetical Protein; Complete proteome.										
SEQUENCE	389 AA; 40230 MW; DB360C24024E93B CRC64;										
QY	9 PAVVPVP 15										
Db	160 PAVVPVP 166										
RESULT	9										
Q9KZK5	PRELIMINARY;	PRM;	416 AA.								
AC	Q9KZK5	PRELIMINARY;	PRM;	436 AA.							
DT	01-OCT-2000 (TREMBLrel. 15, Created)										
DT	01-OCT-2000 (TREMBLrel. 15, Last sequence update)										
DT	01-JUN-2002 (TREMBLrel. 21, Last annotation update)										
DE	Putative secreted protein.										
GN	SG0340 OR SCE34.21C.										
OS	Streptomyces coelicolor.										
OC	Bacteria; Firmicutes; Actinobacteridae; Actinomycetales; Aquificales; Aquificaceae; Aquifex.										
OC	Actinomycetales; Streptomyceinae; Streptomycetaceae; Streptomyces.										
OX	NCBI_TaxID=1902;										
RN	[1]										
RP	SEQUENCE FROM N.A.										
RC	STRAIN=P3(2);										
RC	Cerdano A.M., Parkhill J., Barrell B.G., Rajandream M.A.,										
RL	Submitted (APR-2000) to the EMBL/GenBank/DDBJ databases.										
RN	[3]										
RP	SEQUENCE FROM N.A.										
RC	STRAIN=P3(2);										
RX	MEDLINE=97000351; PubMed=8043436;										
RL	Submitted (APR-2000) to the EMBL/GenBank/DDBJ databases.										
RN	[3]										
RP	SEQUENCE FROM N.A.										
RC	STRAIN=P3(2);										
RN	Medline ID: 97000351; PubMed=8043436;										
RP	SEQUENCE FROM N.A.										
RC	STRAIN=P3(2);										
RN	Medline ID: 97000351; PubMed=8043436;										
RP	SEQUENCE FROM N.A.										
RC	STRAIN=P3(2);										
RN	Medline ID: 97000351; PubMed=8043436;										
RP	SEQUENCE FROM N.A.										
RC	STRAIN=P3(2);										
RN	Medline ID: 97000351; PubMed=8043436;										
RP	SEQUENCE FROM N.A.										
RC	STRAIN=P3(2);										
RN	Medline ID: 97000351; PubMed=8043436;										
RP	SEQUENCE FROM N.A.										
RC	STRAIN=P3(2);										
RN	Medline ID: 97000351; PubMed=8043436;										
RP	SEQUENCE FROM N.A.										
RC	STRAIN=P3(2);										
RN	Medline ID: 97000351; PubMed=8043436;										
RP	SEQUENCE FROM N.A.										
RC	STRAIN=P3(2);										
RN	Medline ID: 97000351; PubMed=8043436;										
RP	SEQUENCE FROM N.A.										
RC	STRAIN=P3(2);										
RN	Medline ID: 97000351; PubMed=8043436;										
RP	SEQUENCE FROM N.A.										
RC	STRAIN=P3(2);										
RN	Medline ID: 97000351; PubMed=8043436;										
RP	SEQUENCE FROM N.A.										
RC	STRAIN=P3(2);										
RN	Medline ID: 97000351; PubMed=8043436;										
RP	SEQUENCE FROM N.A.										
RC	STRAIN=P3(2);										
RN	Medline ID: 97000351; PubMed=8043436;										
RP	SEQUENCE FROM N.A.										
RC	STRAIN=P3(2);										
RN	Medline ID: 97000351; PubMed=8043436;										
RP	SEQUENCE FROM N.A.										
RC	STRAIN=P3(2);										
RN	Medline ID: 97000351; PubMed=8043436;										
RP	SEQUENCE FROM N.A.										
RC	STRAIN=P3(2);										
RN	Medline ID: 97000351; PubMed=8043436;										
RP	SEQUENCE FROM N.A.										
RC	STRAIN=P3(2);										
RN	Medline ID: 97000351; PubMed=8043436;										
RP	SEQUENCE FROM N.A.										
RC	STRAIN=P3(2);										
RN	Medline ID: 97000351; PubMed=8043436;										
RP	SEQUENCE FROM N.A.										
RC	STRAIN=P3(2);										
RN	Medline ID: 97000351; PubMed=8043436;										
RP	SEQUENCE FROM N.A.										
RC	STRAIN=P3(2);										
RN	Medline ID: 97000351; PubMed=8043436;										
RP	SEQUENCE FROM N.A.										
RC	STRAIN=P3(2);										
RN	Medline ID: 97000351; PubMed=8043436;										
RP	SEQUENCE FROM N.A.										
RC	STRAIN=P3(2);										
RN	Medline ID: 97000351; PubMed=8043436;										
RP	SEQUENCE FROM N.A.										
RC	STRAIN=P3(2);										
RN	Medline ID: 97000351; PubMed=8043436;										
RP	SEQUENCE FROM N.A.										
RC	STRAIN=P3(2);										
RN	Medline ID: 97000351; PubMed=8043436;										
RP	SEQUENCE FROM N.A.										
RC	STRAIN=P3(2);										
RN	Medline ID: 97000351; PubMed=8043436;										
RP	SEQUENCE FROM N.A.										
RC	STRAIN=P3(2);										
RN	Medline ID: 97000351; PubMed=8043436;										
RP	SEQUENCE FROM N.A.										
RC	STRAIN=P3(2);										
RN	Medline ID: 97000351; PubMed=8043436;										
RP	SEQUENCE FROM N.A.										
RC	STRAIN=P3(2);										
RN	Medline ID: 97000351; PubMed=8043436;										
RP	SEQUENCE FROM N.A.										
RC	STRAIN=P3(2);										
RN	Medline ID: 97000351; PubMed=8043436;										
RP	SEQUENCE FROM N.A.										
RC	STRAIN=P3(2);										
RN	Medline ID: 97000351; PubMed=8043436;										
RP	SEQUENCE FROM N.A.										
RC	STRAIN=P3(2);										
RN	Medline ID: 97000351; PubMed=8043436;										
RP	SEQUENCE FROM N.A.										
RC	STRAIN=P3(2);										
RN	Medline ID: 97000351; PubMed=8043436;										
RP	SEQUENCE FROM N.A.										
RC	STRAIN=P3(2);										
RN	Medline ID: 97000351; PubMed=8043436;										
RP	SEQUENCE FROM N.A.										
RC	STRAIN=P3(2);										
RN	Medline ID: 97000351; PubMed=8043436;										
RP	SEQUENCE FROM N.A.										
RC	STRAIN=P3(2);										
RN	Medline ID: 97000351; PubMed=8043436;										
RP	SEQUENCE FROM N.A.										
RC	STRAIN=P3(2);										
RN	Medline ID: 97000351; PubMed=8043436;										
RP	SEQUENCE FROM N.A.										
RC	STRAIN=P3(2);										
RN	Medline ID: 97000351; PubMed=8043436;										
RP	SEQUENCE FROM N.A.										
RC	STRAIN=P3(2);										
RN	Medline ID: 97000351; PubMed=8043436;										
RP	SEQUENCE FROM N.A.										
RC	STRAIN=P3(2);										
RN	Medline ID: 97000351; PubMed=8043436;										
RP	SEQUENCE FROM N.A.										
RC	STRAIN=P3(2);										
RN	Medline ID: 97000351; PubMed=8043436;										
RP	SEQUENCE FROM N.A.										
RC	STRAIN=P3(2);										
RN	Medline ID: 97000351; PubMed=8043436;										
RP	SEQUENCE FROM N.A.										
RC	STRAIN=P3(2);										
RN	Medline ID: 97000351; PubMed=8043436;										
RP	SEQUENCE FROM N.A.										
RC	STRAIN=P3(2);										
RN	Medline ID: 97000351; PubMed=8043436;										
RP	SEQUENCE FROM N.A.										
RC	STRAIN=P3(2);										

Query Match 8.9%; Score 7; DB 5; Length 469;  
 Best Local Similarity 100.0%; Pred. No. 48; Mismatches 0;  
 Matches 7; Conservative 0; Indels 0; Gaps 0;

Qy 9 PAAVPVP 15  
 Db 362 PAAVPVP 368

RESULT 12

Q8WP9 8WP9 PRELIMINARY; PRT; 469 AA.  
 AC 08WP9; 01-MAR-2002 (TREMBLrel. 20, Created)  
 DT 01-MAR-2002 (TREMBLrel. 20, last sequence update)  
 DE zeste protein (Fragment).  
 GN ZESTE.  
 OS Drosophila melanogaster (Fruit fly).  
 Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephyoidea; Drosophilidae; Drosophila.  
 NCBI\_TaxID=7227;  
 RN [1] SEQUENCE FROM N.A.  
 RP STRAIN=K2, K19, AND K22;  
 RA Zangerl, B.;  
 RT "Evidence for selection in a natural population of drosophila  
 melanogaster.";  
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AJ399788; CAC01395\_1; -  
 DR EMBL; AJ399795; CAC01402\_1; -  
 DR EMBL; AJ399798; CAC01405\_1; -  
 FT NON\_TER 1  
 SQ SEQUENCE 469 AA; 50996 MW; 474BF9F575D4DC24 CRC64;

Query Match 8.9%; score 7; DB 5; Length 469;  
 Best Local Similarity 100.0%; Pred. No. 48;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 9 PAAVPVP 15  
 Db 362 PAAVPVP 368

RESULT 13

Q8WP54 8WP54 PRELIMINARY; PRT; 469 AA.  
 AC 08WP54; 01-MAR-2002 (TREMBLrel. 20, Created)  
 DT 01-MAR-2002 (TREMBLrel. 20, last sequence update)  
 DE zeste protein (Fragment).  
 GN ZESTE.  
 OS Drosophila melanogaster (Fruit fly).  
 Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephyoidea; Drosophilidae; Drosophila.  
 NCBI\_TaxID=7227;  
 RN [1] SEQUENCE FROM N.A.  
 RP STRAIN=ND5, ND7, ND11, ND13, AND ND19;  
 RA Zangerl, B.;  
 RT "Evidence for selection in a natural population of Drosophila  
 melanogaster.", submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL; AJ399747; CAC01411\_1; -  
 DR EMBL; AJ399748; CAC01412\_1; -  
 DR EMBL; AJ399749; CAC01413\_1; -  
 DR EMBL; AJ399751; CAC01415\_1; -  
 DR EMBL; AJ399756; CAC01421\_1; -

RESULT 14

Q8WQ65 8WQ65 PRELIMINARY; PRT; 471 AA.  
 AC 08WQ65; 01-MAR-2002 (TREMBLrel. 20, Created)  
 DT 01-MAR-2002 (TREMBLrel. 20, last sequence update)  
 DE zeste protein (Fragment).  
 GN ZESTE.  
 OS Drosophila melanogaster (Fruit fly).  
 Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephyoidea; Drosophilidae; Drosophila.  
 NCBI\_TaxID=7227;  
 RN [1] SEQUENCE FROM N.A.  
 RP STRAIN=K8;  
 RA Zangerl, B.;  
 RT "Evidence for selection in a natural population of Drosophila  
 melanogaster.";  
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AJ399789; CAC01396\_1; -  
 FT NON\_TER 1  
 SQ SEQUENCE 471 AA; 51194 MW; 07E4221753E96B20 CRC64;

Query Match 8.9%; Score 7; DB 5; Length 471;  
 Best Local Similarity 100.0%; Pred. No. 48;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 9 PAAVPVP 15  
 Db 364 PAAVPVP 370

RESULT 15

Q8WQ64 8WQ64 PRELIMINARY; PRT; 471 AA.  
 AC 08WQ64; 01-MAR-2002 (TREMBLrel. 20, Created)  
 DT 01-MAR-2002 (TREMBLrel. 20, last sequence update)  
 DE 01-MAR-2002 (TREMBLrel. 20, last annotation update)  
 DE zeste protein (Fragment).  
 GN ZESTE.  
 OS Drosophila melanogaster (Fruit fly).  
 Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephyoidea; Drosophilidae; Drosophila.  
 NCBI\_TaxID=7227;  
 RN [1] SEQUENCE FROM N.A.  
 RP STRAIN=K11;  
 RA Zangerl, B.;  
 RT "Evidence for selection in a natural population of Drosophila  
 melanogaster.";  
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AJ399791; CAC01398\_1; -  
 FT NON\_TER 1  
 SQ SEQUENCE 471 AA; 51199 MW; 075DB4C342F97CE4 CRC64;

Wed Jan 15 08:39:02 2003

us-09-506-079e-1.olig.rspt

Page 6

Query Match 8.9%; Score 7; DB 5; Length 471;  
Best Local Similarity 100.0%; Pred. No. 48;  
Matches 7; Conservative 0; Mismatches 0;  
Oy 9 PAAPVVP 15  
Db 364 PAAPVVP 370  
|||||||  
0; Indels 0; Gaps 0;

Search completed: January 14, 2003, 17:14:22  
Job time : 11.7249 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

### OM protein - protein search, using sw model

Run on: January 14, 2003, 17:14:40 ; search time 2.69679 seconds

(without alignments)  
568.337 million cell updates/sec

Title:	US-09-506-079E-1	20	6	7.6	382	9	US-09-941-947A-30
Perfect score:	79	21	6	7.6	398	10	US-09-741-664-430
Sequence:	1 GKHSSXPRPAAVPVPARQP . . . . . VGRGXDPDAHVAVLSRYEG 79	22	6	7.6	404	10	US-09-764-864-2244
Scoring table:	OLIGO	23	6	7.6	408	9	US-10-027-806-744
Gapop:	60.0 , Gapext 60.0	24	6	7.6	408	9	US-10-034-623-74
Searched:	118974 seqs, 1941057 residues	25	6	7.6	418	10	US-09-815-242-2003
Word size :	0	26	6	7.6	450	10	US-09-764-864-804
Total number of hits satisfying chosen parameters:	118974	27	6	7.6	685	9	US-10-029-180-82
Minimum DB seq length: 0		28	6	7.6	720	10	US-09-919-497-83
Maximum DB seq length: 200000000		29	6	7.6	784	10	US-09-770-595A-3

### Post-processing: Listing first 45 summaries

Database : Published\_Applications\_AA:\*

```

1: /cgn2_6/ptoda/2/pubpaas/PCT_NEW_PUB_pep:*
2: /cgn2_6/ptoda/2/pubpaas/PCT_NEW_PUB_pep:*
3: /cgn2_6/ptoda/2/pubpaas/US06_NEWPUBCOMB_pep:*
4: /cgn2_6/ptoda/2/pubpaas/US06_PUBCOMB_pep:*
5: /cgn2_6/ptoda/2/pubpaas/US07_NEWPUBCOMB_pep:*
6: /cgn2_6/ptoda/2/pubpaas/US07_PUBCOMB_pep:*
7: /cgn2_6/ptoda/2/pubpaas/PCTUS08_PUBCOMB_pep:*
8: /cgn2_6/ptoda/2/pubpaas/US08_PUBCOMB_pep:*
9: /cgn2_6/ptoda/2/pubpaas/US09_NEWPUBCOMB_pep:*
10: /cgn2_6/ptoda/2/pubpaas/US09_PUBCOMB_pep:*
11: /cgn2_6/ptoda/2/pubpaas/US10_NEWPUBCOMB_pep:*
12: /cgn2_6/ptoda/2/pubpaas/US10_PUBCOMB_pep:*
13: /cgn2_6/ptoda/2/pubpaas/US10_NEWPUBCOMB_pep:*
14: /cgn2_6/ptoda/2/pubpaas/US60_PUBCOMB_pep:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

### SUMMARIES

Result No.	Scored Match Length	Query DB ID	Description	ALIGNMENTS
1	7	8.9	61 10 US-09-864-761-44740 Sequence 44140, A	Sequence 30, Appl
2	7	8.9	215 9 US-10-001-876-210 Sequence 210, App	Sequence 430, App
3	6	7.6	38 10 US-09-864-761-41933 Sequence 41933, A	Sequence 1244, App
4	6	7.6	68 10 US-09-764-887-226 Sequence 226, App	Sequence 74, Appl
5	6	7.6	86 10 US-09-869-869-1118 Sequence 1118, App	Sequence 74, Appl
6	6	7.6	168 10 US-09-925-301-1470 Sequence 1470, App	Sequence 12003, A
7	6	7.6	238 12 US-10-024-579-16 Sequence 16, App	Sequence 804, App
8	6	7.6	257 12 US-10-024-579-14 Sequence 14, App	Sequence 82, Appl
9	6	7.6	264 12 US-10-024-579-12 Sequence 12, App	Sequence 83, Appl
10	6	7.6	267 9 US-09-808-602-23 Sequence 23, App	Sequence 3, Appl
11	6	7.6	9 US-09-808-602-25 Sequence 25, App	Sequence 75, Appl
12	6	7.6	283 12 US-10-024-579-10 Sequence 10, App	Sequence 5338, App
13	6	7.6	290 9 US-10-041-006A-7 Sequence 7, Appl	Sequence 1223, A
14	6	7.6	291 9 US-10-028-1222 Sequence 222, App	Sequence 347, App
15	6	7.6	290 12 US-10-040-655-7 Sequence 7, Appl	Sequence 2593, App
16	6	7.6	298 9 US-09-764-853-552 Sequence 552, App	Sequence 303, App
17	6	7.6	299 9 US-09-808-602-63 Sequence 63, Appl	Sequence 40862, App
18	6	7.6	323 10 US-09-764-864-1265 Sequence 1255, App	Sequence 41246, A
19	6	7.6	349 10 US-09-766-366-4 Sequence 4, Appl	Sequence 43417, A
				Sequence 39863, A
				Sequence 36827, A

RESULT 1

US-09-864-761-44740

; Sequence 44140, Application US/09864761

; Patent No. US20030048763A1

; GENERAL INFORMATION:

; APPLICANT: Penn, Sharron G.

; APPLICANT: Rank, David R.

; APPLICANT: Hanzel, David R.

; TITLE OF INVENTION: Chen, Weisheng

; TITLE OF INVENTION: GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO

; FILE REFERENCE: Neomica-X.1

; CURRENT APPLICATION NUMBER: US/09-864,761

; CURRENT FILING DATE: 2001-05-23

; PRIOR APPLICATION NUMBER: US 60/207,456

; PRIOR FILING DATE: 2000-02-04

; PRIOR APPLICATION NUMBER: US 09/632,366

; PRIOR FILING DATE: 2000-08-03

; PRIOR APPLICATION NUMBER: GB 24263,5

; PRIOR APPLICATION NUMBER: US 60/236,359

; PRIOR FILING DATE: 2000-05-27

; PRIOR APPLICATION NUMBER: PCT/US01/00666

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00667

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00664

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00669

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00665

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00668

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00663

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00662

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00661

; PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00670  
 PRIOR FILING DATE: 2001-01-30  
 PRIOR APPLICATION NUMBER: US 60/234,687  
 PRIOR FILING DATE: 2000-09-21  
 PRIOR APPLICATION NUMBER: US 09/608,408  
 PRIOR FILING DATE: 2000-06-30  
 PRIOR APPLICATION NUMBER: US 09/774,203  
 PRIOR FILING DATE: 2001-01-29  
 NUMBER OF SEQ ID NOS.: 49117  
 SOFTWARE: Annotax Sequence Listing Engine vers. 1.1  
 SEQ ID NO 44740  
 LENGTH: 61  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 FEATURE:  
 OTHER INFORMATION: MAP TO AC016057.3  
 OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.72  
 OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.59  
 OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.53  
 OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.52  
 OTHER INFORMATION: SWISSPROT HIT: Q13563, EVALU 2.00e+00  
 OTHER INFORMATION: EST\_HUMAN HIT: BP570694.1, EVALU 4.00e-16  
 US-09-864-761-44740

Query Match 8.9%; Score 7; DB 10; Length 61;  
 Best Local Similarity 100.0%; Pred. 2.2%; Mismatches 0; Indels 0; Gaps 0;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 23 AHPVLSF 29  
 Db 37 AHPVLSF 43

RESULT 2  
 US-10-001-876-210  
 ; Sequence 210, Application US/10001876  
 ; Patent No. US200201771401  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Salceda, Susana  
 ; APPLICANT: Macina, Roberto  
 ; APPLICANT: Recipon, Helve  
 ; APPLICANT: Cafferkey, Robert  
 ; APPLICANT: Ali, Shujath  
 ; APPLICANT: Sun, Yongming  
 ; APPLICANT: Liu, Chenghua  
 ; TITLE OF INVENTION: Compositions and Methods Relating to Prostate Specific Genes and  
 ; CURRENT APPLICATION NUMBER: US/10/001,876  
 ; FILE REFERENCE: DE1-0285  
 ; CURRENT FILING DATE: 2001-11-20  
 ; PRIOR APPLICATION NUMBER: US/10/001,876  
 ; NUMBER OF SEQ ID NOS: 211  
 ; SOFTWARE: Patentin version 3.1  
 ; SEQ ID NO 210  
 ; LENGTH: 215  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapien

US-10-001-876-210  
 ; Query Match 8.9%; Score 7; DB 9; Length 215;  
 ; Best Local Similarity 100.0%; Pred. No. 7; Mismatches 0; Indels 0; Gaps 0;  
 ; Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 27 LSFLRPS 33  
 Db 60 LSFLRPS 66

RESULT 3  
 US-09-864-761-41953  
 ; Sequence 41953, Application US/09864761  
 ; Patent No. US2002048763A1  
 ; GENERAL INFORMATION:

Query Match 7.6%; Score 6; DB 10; Length 38;  
 Best Local Similarity 100.0%; Pred. No. 13; Mismatches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 47 PLSPTS 52  
 Db 24 PLSPTS 29

RESULT 4

US-09-764-887-226  
; Sequence 226 Application US/09764887  
; Patent No. US2002042096A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PA113  
; CURRENT APPLICATION NUMBER: US/09/764,887  
; CURRENT FILING DATE: 2001-01-17  
; PRIORITY APPLICATION NUMBER: US 2000-03-08  
; PRIORITY FILING DATE: 1999-03-12  
; NUMBER OF SEQ ID NOS: 658  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO: 226  
; LENGTH: 68  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE: SITE  
; LOCATION: (6)  
; OTHER INFORMATION: xaa equals any of the naturally occurring L-amino acids  
; Query Match 7.6%; Score 6; DB 10; Length 68;  
; Best Local Similarity 100.0%; Pred. No. 23;  
; Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
; QY 55 ISPVSV 60  
; Db 28 ISPVSV 33

RESULT 5  
US-09-764-869-1118  
; Sequence 1118, Application US/09764869  
; Patent No. US2002061521A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PC007  
; CURRENT APPLICATION NUMBER: US/09/764,869  
; CURRENT FILING DATE: 2001-01-17  
; PRIORITY APPLICATION DATA REMOVED - refer to PALM or file wrapper  
; NUMBER OF SEQ ID NOS: 2442  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO: 1118  
; LENGTH: 86  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE: SITE  
; LOCATION: (4)  
; OTHER INFORMATION: xaa equals any of the naturally occurring L-amino acids  
; NAME/KEY: SITE  
; LOCATION: (48)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
; US-09-764-869-1118

Query Match 7.6%; Score 6; DB 10; Length 86;  
; Best Local Similarity 100.0%; Pred. No. 29;  
; Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
; QY 43 LPLAPL 48  
; Db 61 LPLAPL 66

RESULT 6  
US-09-925-301-1470  
; Sequence 1470, Application US/09925301  
; Patent No. US2002052308A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
; FILE REFERENCE: PA113  
; CURRENT APPLICATION NUMBER: US/09/925,301  
; CURRENT FILING DATE: 2001-12-18  
; PRIORITY APPLICATION NUMBER: US 60/258,595  
; PRIORITY FILING DATE: 2000-12-28  
; NUMBER OF SEQ ID NOS: 17  
; SOFTWARE: FastSEQ for Windows Version 4.0

FILE REFERENCE: PA105  
; CURRENT APPLICATION NUMBER: US/09/925,301  
; CURRENT FILING DATE: 2001-08-10  
; PRIORITY APPLICATION NUMBER: PCT/US00/05882  
; PRIORITY FILING DATE: 2000-03-08  
; NUMBER OF SEQ ID NOS: 1694  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO: 1470  
; LENGTH: 168  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE: SITE  
; LOCATION: (17)  
; OTHER INFORMATION: xaa equals any of the naturally occurring L-amino acids  
; NAME/KEY: SITE  
; LOCATION: (136)  
; OTHER INFORMATION: xaa equals any of the naturally occurring L-amino acids  
; NAME/KEY: SITE  
; LOCATION: (139)  
; OTHER INFORMATION: xaa equals any of the naturally occurring L-amino acids  
; NAME/KEY: SITE  
; LOCATION: (141)  
; OTHER INFORMATION: xaa equals any of the naturally occurring L-amino acids  
; NAME/KEY: SITE  
; LOCATION: (143)  
; OTHER INFORMATION: xaa equals any of the naturally occurring L-amino acids  
; NAME/KEY: SITE  
; LOCATION: (146)  
; OTHER INFORMATION: xaa equals any of the naturally occurring L-amino acids  
; NAME/KEY: SITE  
; LOCATION: (148)  
; OTHER INFORMATION: xaa equals any of the naturally occurring L-amino acids  
; NAME/KEY: SITE  
; LOCATION: (152)  
; OTHER INFORMATION: xaa equals any of the naturally occurring L-amino acids  
; NAME/KEY: SITE  
; LOCATION: (153)  
; OTHER INFORMATION: xaa equals any of the naturally occurring L-amino acids  
; NAME/KEY: SITE  
; LOCATION: (158)  
; OTHER INFORMATION: xaa equals any of the naturally occurring L-amino acids  
; US-09-925-301-1470

Query Match 7.6%; Score 6; DB 10; Length 168;  
; Best Local Similarity 100.0%; Pred. No. 54;  
; Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
; QY 25 PVLSFL 30  
; Db 121 PVLSFL 126

RESULT 7  
US-10-024-579-16  
; Sequence 16, Application US/10024579  
; Patent No. US20020119522A1  
; GENERAL INFORMATION:  
; APPLICANT: Friddle, Carl Johan  
; APPLICANT: Gerhardt, Brend  
; APPLICANT: Hilbun, Erin  
; APPLICANT: Turner, C. Alexander Jr.  
; TITLE OF INVENTION: No. US20020119522A1 Human Ion Channel-Related Proteins  
; TITLE OF INVENTION: And Polynucleotides Encoding the Same  
; FILE REFERENCE: LEX-0774-USA  
; CURRENT APPLICATION NUMBER: US/10/024,579  
; CURRENT FILING DATE: 2001-12-18  
; PRIORITY APPLICATION NUMBER: US 60/258,595  
; PRIORITY FILING DATE: 2000-12-28  
; NUMBER OF SEQ ID NOS: 17  
; SOFTWARE: FastSEQ for Windows Version 4.0

Query Match Best Local Similarity 100.0%; Score 6; DB 12; Length 238; Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0; QY 26 VLSFLR 31  
Db 87 VLSFLR 92

RESULT 8 US-10-024-579-14 Sequence 14, Application US/10024579  
; Patent No. US20020119522A1  
GENERAL INFORMATION:  
; APPLICANT: Fiddle, Carl Johan  
; APPLICANT: Gerhardt, Brenda  
; APPLICANT: Hilbun, Erin  
; APPLICANT: Turner, C. Alexander Jr.  
TITLE OF INVENTION: Human Ion Channel-Related Proteins and Polynucleotides Encoding the Same  
FILE REFERENCE: LEX-0274-USA  
CURRENT APPLICATION NUMBER: US/10/024,579  
CURRENT FILING DATE: 2001-12-18  
PRIOR APPLICATION NUMBER: US 60/258,595  
PRIOR FILING DATE: 2000-12-28  
NUMBER OF SEQ ID NOS: 17  
SOFTWARE: FastSEQ for Windows Version 4.0  
SEQ ID NO 14  
LENGTH: 257  
TYPE: PRT  
ORGANISM: homo sapiens  
US-10-024-579-14

Query Match Best Local Similarity 100.0%; Score 6; DB 12; Length 257; Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0; QY 26 VLSFLR 31  
Db 87 VLSFLR 92

RESULT 9 US-10-024-579-12 Sequence 12, Application US/10024579  
; Patent No. US20020119522A1  
GENERAL INFORMATION:  
; APPLICANT: Fiddle, Carl Johan  
; APPLICANT: Gerhardt, Brenda  
; APPLICANT: Hilbun, Erin  
; APPLICANT: Turner, C. Alexander Jr.  
TITLE OF INVENTION: Human Ion Channel-Related Proteins and Polynucleotides Encoding the Same  
FILE REFERENCE: LEX-0274-USA  
CURRENT APPLICATION NUMBER: US/10/024,579  
CURRENT FILING DATE: 2001-12-18  
PRIOR APPLICATION NUMBER: US 60/258,595  
PRIOR FILING DATE: 2000-12-28  
NUMBER OF SEQ ID NOS: 17  
SOFTWARE: FastSEQ for Windows Version 4.0  
SEQ ID NO 12  
LENGTH: 264  
TYPE: PRT  
ORGANISM: homo sapiens  
US-10-024-579-12

RESULT 10 US-03-808-602-23 Sequence 23, Application US/09808602  
; Patent No. US20020155115A1  
GENERAL INFORMATION:  
; APPLICANT: Vernet, Corinne A  
; APPLICANT: Fernandes, Elma  
; APPLICANT: Herrman, John L  
; APPLICANT: Majumder, Kumud  
; APPLICANT: Mishra, Vishnu  
; APPLICANT: Mezes, Peter S  
; APPLICANT: Macdougall, John  
; APPLICANT: Fernandes, Elma  
; APPLICANT: Shimkets, Richard A  
; APPLICANT: Herrman, John L  
; APPLICANT: Majumder, Kumud  
; APPLICANT: Mishra, Vishnu  
; APPLICANT: Mezes, Peter S  
; APPLICANT: Macdougall, John  
TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same  
FILE REFERENCE: 15966-697 CIP  
CURRENT APPLICATION NUMBER: US/09/808,602  
CURRENT FILING DATE: 2001-03-05  
PRIOR APPLICATION NUMBER: 60/186,596  
PRIOR FILING DATE: 2000-03-03  
NUMBER OF SEQ ID NOS: 114  
SOFTWARE: Patentin Ver. 2.1  
SEQ ID NO 23  
LENGTH: 267  
TYPE: PRT  
ORGANISM: homo sapiens  
US-03-808-602-23

RESULT 11 US-03-808-602-25 Sequence 25, Application US/09808602  
; Patent No. US20020155115A1  
GENERAL INFORMATION:  
; APPLICANT: Vernet, Corinne A  
; APPLICANT: Fernandes, Elma  
; APPLICANT: Shinkets, Richard A  
; APPLICANT: Herrman, John L  
; APPLICANT: Majumder, Kumud  
; APPLICANT: Mishra, Vishnu  
; APPLICANT: Mezes, Peter S  
; APPLICANT: Macdougall, John  
TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same  
FILE REFERENCE: 15966-697 CIP  
CURRENT APPLICATION NUMBER: US/09/808,602  
CURRENT FILING DATE: 2001-03-14  
PRIOR APPLICATION NUMBER: 60/186,596  
PRIOR FILING DATE: 2000-03-05  
PRIOR APPLICATION NUMBER: 60/186,596  
PRIOR FILING DATE: 2000-03-03  
NUMBER OF SEQ ID NOS: 114  
SOFTWARE: Patentin Ver. 2.1  
SEQ ID NO 25  
LENGTH: 267  
TYPE: PRT

US-09-808-602-25

RESULT 12

Query Match 7.6%; Score 6; DB 9; Length 267;

Best Local Similarity 100.0%; Pred. No. 83; Mismatches 0; Indels 0; Gaps 0;

Qy 43 LPALPL 48

Db 20 LPALPL 25

RESULT 14

US-10-028-072-222

Sequence 222, Application US/10028072

Publication No. US20030004311A1

GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.

APPLICANT: Bersini, Maureen

APPLICANT: DeForge, Laura

APPLICANT: Desnoyers, Luc

APPLICANT: Flvaroff, Ellen

APPLICANT: Gao, Wei-Qiang

APPLICANT: Gerritsen, Mary E.

APPLICANT: Goddard, Audrey

APPLICANT: Godowski, Paul J.

APPLICANT: Gurney, Austin L.

APPLICANT: Shewwood, Steven

APPLICANT: Smith, Victoria

APPLICANT: Stewart, Timothy A.

APPLICANT: Tunas, Daniel

APPLICANT: Watanabe, Colin K

APPLICANT: Wood, William

APPLICANT: Zhang

TITLE OF INVENTION:

FILE REFERENCE:

CURRENT APPLICATION NUMBER: US/10/028,072

CURRENT FILING DATE: 2001-12-19

PRIOR APPLICATION NUMBER: 60/049911

PRIOR FILING DATE: 1997-05-18

PRIOR APPLICATION NUMBER: 60/056974

PRIOR FILING DATE: 1997-05-26

PRIOR APPLICATION NUMBER: 60/059113

PRIOR FILING DATE: 1997-05-17

PRIOR APPLICATION NUMBER: 60/059115

PRIOR FILING DATE: 1997-05-17

PRIOR APPLICATION NUMBER: 60/059117

PRIOR FILING DATE: 1997-05-17

PRIOR APPLICATION NUMBER: 60/059122

PRIOR FILING DATE: 1997-05-17

PRIOR APPLICATION NUMBER: 60/059184

PRIOR FILING DATE: 1997-05-17

PRIOR APPLICATION NUMBER: 60/059263

PRIOR FILING DATE: 1997-05-18

PRIOR APPLICATION NUMBER: 60/059352

PRIOR FILING DATE: 1997-05-19

PRIOR APPLICATION NUMBER: 60/059588

PRIOR FILING DATE: 1997-05-19

PRIOR APPLICATION NUMBER: 60/059836

PRIOR FILING DATE: 1997-05-24

PRIOR APPLICATION NUMBER: 60/062250

PRIOR FILING DATE: 1997-10-17

PRIOR APPLICATION NUMBER: 60/062285

PRIOR FILING DATE: 1997-10-17

PRIOR APPLICATION NUMBER: 60/062287

PRIOR FILING DATE: 1997-10-17

PRIOR APPLICATION NUMBER: 60/062814

PRIOR FILING DATE: 1997-10-24

PRIOR APPLICATION NUMBER: 60/062816

PRIOR FILING DATE: 1997-10-24

PRIOR APPLICATION NUMBER: 60/063045

PRIOR FILING DATE: 1997-10-24

PRIOR APPLICATION NUMBER: 60/063082

PRIOR FILING DATE: 1997-10-31

PRIOR APPLICATION NUMBER: 60/063127

PRIOR FILING DATE: 1997-10-24

PRIOR APPLICATION NUMBER: 60/063327

PRIOR FILING DATE: 1997-10-27

PRIOR APPLICATION NUMBER: 60/063329

PRIOR FILING DATE: 1997-10-27

PRIOR APPLICATION NUMBER: 60/063550

PRIOR FILING DATE: 1997-10-28

PRIOR APPLICATION NUMBER: 60/063561

PRIOR FILING DATE: 1997-10-28

US-10-024-579-10

RESULT 13

Query Match 7.6%; Score 6; DB 9; Length 290;

Best Local Similarity 100.0%; Pred. No. 89; Mismatches 0; Indels 0; Gaps 0;

Qy 8 RPAAVP 13

Db 3 RPAAVP 8

PRIOR APPLICATION NUMBER: 60/063704  
 PRIOR FILING DATE: 1997-10-29  
 PRIOR APPLICATION NUMBER: 60/063733  
 PRIOR FILING DATE: 1997-10-29  
 PRIOR APPLICATION NUMBER: 60/063735  
 PRIOR FILING DATE: 1997-10-29  
 PRIOR APPLICATION NUMBER: 60/063738  
 PRIOR FILING DATE: 1997-10-29  
 PRIOR APPLICATION NUMBER: 60/063755  
 PRIOR FILING DATE: 1997-10-17  
 PRIOR APPLICATION NUMBER: 60/064248  
 PRIOR FILING DATE: 1997-11-03  
 PRIOR APPLICATION NUMBER: 60/064809  
 PRIOR FILING DATE: 1997-11-07  
 PRIOR APPLICATION NUMBER: 60/065186  
 PRIOR FILING DATE: 1997-11-12  
 PRIOR APPLICATION NUMBER: 60/065846  
 PRIOR FILING DATE: 1997-11-17  
 PRIOR APPLICATION NUMBER: 60/066364  
 PRIOR FILING DATE: 1997-11-21  
 PRIOR APPLICATION NUMBER: 60/066453  
 PRIOR FILING DATE: 1997-11-24  
 PRIOR APPLICATION NUMBER: 60/066511  
 PRIOR FILING DATE: 1997-11-24  
 PRIOR APPLICATION NUMBER: 60/066770  
 PRIOR FILING DATE: 1997-11-24  
 PRIOR APPLICATION NUMBER: 60/069212  
 PRIOR FILING DATE: 1997-12-11  
 PRIOR APPLICATION NUMBER: 60/069278  
 PRIOR FILING DATE: 1997-12-11  
 PRIOR APPLICATION NUMBER: 60/069334  
 PRIOR FILING DATE: 1997-12-11  
 PRIOR APPLICATION NUMBER: 60/069594  
 PRIOR FILING DATE: 1997-12-16  
 PRIOR APPLICATION NUMBER: 60/072320  
 PRIOR FILING DATE: 1998-01-23  
 PRIOR APPLICATION NUMBER: 60/073612  
 PRIOR FILING DATE: 1998-02-04  
 PRIOR APPLICATION NUMBER: 60/074086  
 PRIOR FILING DATE: 1998-02-09  
 PRIOR APPLICATION NUMBER: 60/074092  
 PRIOR FILING DATE: 1998-02-09  
 PRIOR APPLICATION NUMBER: 60/077791  
 PRIOR FILING DATE: 1998-03-12  
 PRIOR APPLICATION NUMBER: 60/078910  
 PRIOR FILING DATE: 1998-03-20  
 PRIOR APPLICATION NUMBER: 60/079294  
 PRIOR FILING DATE: 1998-03-25  
 PRIOR APPLICATION NUMBER: 60/079663  
 PRIOR FILING DATE: 1998-02-27  
 PRIOR APPLICATION NUMBER: 60/079728  
 PRIOR FILING DATE: 1998-03-27  
 PRIOR APPLICATION NUMBER: 60/080165  
 PRIOR FILING DATE: 1998-03-31  
 PRIOR APPLICATION NUMBER: 60/081203  
 PRIOR FILING DATE: 1998-04-09  
 PRIOR APPLICATION NUMBER: 60/081229  
 PRIOR FILING DATE: 1998-04-09  
 PRIOR APPLICATION NUMBER: 60/081695  
 PRIOR FILING DATE: 1998-04-14  
 PRIOR APPLICATION NUMBER: 60/081817  
 PRIOR FILING DATE: 1998-04-15  
 PRIOR APPLICATION NUMBER: 60/081818  
 PRIOR FILING DATE: 1998-04-15  
 PRIOR APPLICATION NUMBER: 60/082999  
 PRIOR FILING DATE: 1998-04-24  
 PRIOR APPLICATION NUMBER: 60/083322  
 PRIOR FILING DATE: 1998-04-28  
 PRIOR APPLICATION NUMBER: 60/083545  
 PRIOR FILING DATE: 1998-04-29  
 PRIOR APPLICATION NUMBER: 60/084600  
 PRIOR FILING DATE: 1998-05-07  
 PRIOR APPLICATION NUMBER: 60/084627

PRIOR FILING DATE: 1998-05-13  
 PRIOR APPLICATION NUMBER: 60/085338  
 PRIOR FILING DATE: 1998-05-13  
 PRIOR APPLICATION NUMBER: 60/085339  
 PRIOR FILING DATE: 1998-05-13  
 PRIOR APPLICATION NUMBER: 60/085523  
 PRIOR FILING DATE: 1998-05-15  
 PRIOR APPLICATION NUMBER: 60/085697  
 PRIOR FILING DATE: 1998-05-15  
 PRIOR APPLICATION NUMBER: 60/085704  
 PRIOR FILING DATE: 1998-05-15  
 PRIOR APPLICATION NUMBER: 60/086414  
 PRIOR FILING DATE: 1998-05-22  
 PRIOR APPLICATION NUMBER: 60/086430  
 PRIOR FILING DATE: 1998-05-22  
 PRIOR APPLICATION NUMBER: 60/087106  
 PRIOR FILING DATE: 1998-05-28  
 PRIOR APPLICATION NUMBER: 60/088026  
 PRIOR FILING DATE: 1998-06-04  
 PRIOR APPLICATION NUMBER: 60/088730  
 PRIOR FILING DATE: 1998-06-10  
 PRIOR APPLICATION NUMBER: 60/088741  
 PRIOR FILING DATE: 1998-06-10  
 PRIOR APPLICATION NUMBER: 60/088810  
 PRIOR FILING DATE: 1998-06-10  
 PRIOR APPLICATION NUMBER: 60/088858  
 PRIOR FILING DATE: 1998-06-11  
 PRIOR APPLICATION NUMBER: 60/089532  
 PRIOR FILING DATE: 1998-06-17  
 PRIOR APPLICATION NUMBER: 60/089599  
 PRIOR FILING DATE: 1998-06-17  
 PRIOR APPLICATION NUMBER: 60/089907  
 PRIOR FILING DATE: 1998-06-18  
 PRIOR APPLICATION NUMBER: 60/089947  
 PRIOR FILING DATE: 1998-06-19  
 PRIOR APPLICATION NUMBER: 60/090349  
 PRIOR FILING DATE: 1998-06-23  
 PRIOR APPLICATION NUMBER: 60/090429  
 PRIOR FILING DATE: 1998-06-24  
 PRIOR APPLICATION NUMBER: 60/090445  
 PRIOR FILING DATE: 1998-06-24  
 PRIOR APPLICATION NUMBER: 60/090538  
 PRIOR FILING DATE: 1998-06-24  
 PRIOR APPLICATION NUMBER: 60/090863  
 PRIOR FILING DATE: 1998-06-26  
 PRIOR APPLICATION NUMBER: 60/091360  
 PRIOR FILING DATE: 1998-07-01  
 PRIOR APPLICATION NUMBER: 60/091519  
 PRIOR FILING DATE: 1998-07-02  
 PRIOR APPLICATION NUMBER: 60/091982  
 PRIOR FILING DATE: 1998-07-07

Query Match      7.6%; Score 6; DB 9; Length 290;  
 Best Local Similarity 100.0%; Pred. No. 89;  
 Matches 6; Conservative 0; Mismatches 0;  
 Indels 0; Gaps 0;

Qy    8 RPAAVP 13  
 Db    3 RPAAVP 8

RESULT 15

US-10-04-655-7  
 ; Sequence 7, Application US/10040655  
 ; Patent No. US200246802A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Andrade-Gordon, Patricia

APPLICANT: Darrow, Andrew  
APPLICANT: Q1, Jian-shen  
TITLE OF INVENTION: DNA encoding the novel human serine  
FILE REFERENCE: protease\_T  
CURRENT APPLICATION NUMBER: US10/040,655  
CURRENT FILING DATE: 2002-01-07  
NUMBER OF SEQ ID NOS: 11  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO: 7  
LENGTH: 290  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-040-655-7

Query Match 7.6%; Score 6; DB 12; Length 290;  
Best Local Similarity 100.0%; Pred No. 89;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 8 RPAAVP 13  
| | | |  
Db 3 RPAAVP 8

Search completed: January 14, 2003, 17:21:49  
Job time : 4.69679 secs

THIS PAGE BLANK (USPTO)

Gencore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

## OM protein - protein search, using sw mode!

Run on: January 14, 2003, 17:10:35 ; Search time 2.69679 Seconds  
(without alignments)

1215.012 Million cell updates/sec

Title: US-09-506-079E-1  
perfect score: 79  
Sequence: 1 GHXNSXPRPAAVPVPRXQP . . . . . VGRGXDPDAHVAVXLRSYEG 79

Scoring table: OLIGO Gapop 60.0 , Gapext 60.0

Searched: 112892 seqs, 41476328 residues

Word size : 0

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : Swissprot\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	% Score	Query Match Length	DB ID	Description
1	7	8.9	114	GON2_TUPGB
2	7	8.9	380	YMH7_CAEEL
3	7	8.9	574	ZEST_DROME
4	7	8.9	635	SL56_HUMAN
5	7	8.9	677	GCP3_MOUSE
6	7	8.9	907	GCP3_HUMAN
7	6	56	1	PSBK_PINTH
8	6	7.6	126	MFI4_MAIZE
9	6	7.6	145	VP3_BCPHP
10	6	7.6	195	GSPJ_ECOLI
11	6	7.6	198	LMBV_CHICK
12	6	7.6	204	LAF4_MOUSE
13	6	7.6	206	COX3_BACFT
14	6	7.6	207	YQQP_BACSU
15	6	7.6	209	ADPP_ECOLI
16	6	7.6	215	CYB6_ODOST
17	6	7.6	220	ACPD_STRCC
18	6	7.6	220	B1OD_AQUAE
19	6	7.6	239	PRR4_BOVIN
20	6	7.6	241	LAT_RAT
21	6	7.6	242	1_LAT_MOUSE
22	6	7.6	248	PT16_LYCES
23	6	7.6	260	1_XH2_PASMU
24	6	7.6	263	COQ4_CAELF
25	6	7.6	269	1_NIFP_AZOCH
26	6	7.6	290	1_AR1_MESAU
27	6	7.6	290	1_MPN_HUMAN
28	6	7.6	301	PTY34_SYN3
29	6	7.6	313	1_XKB_ECOLI
30	6	7.6	313	1_Y135_TREPA
31	6	7.6	322	1_CYSL_HOMAM
32	6	7.6	344	1_LPKX_NEIMA
33	6	7.6	344	1_QJY4_neisseria_m
34				1_LPXB_NEIMB

## ALIGNMENTS

RESULT	ID	NAME	STANDARD	PRT:	AA:
1	GON2_TUPGB				
	ID: GON2_TUPGB				
	RT: 095336;				
	DT: 15-DEC-1998	(Rel. 37, Created)			
	DT: 15-DEC-1998	(Rel. 37, last sequence update)			
	DT: 30-MAY-2000	(Rel. 39, last annotation update)			
	DE: Propranololiberin II precursor	[Contains: Gonadoliberin II (LHRH II) (Luteinizing hormone releasing hormone II) (Gonadotropin releasing hormone II) (GnRH II); GnRH-associated peptide III; GNRH2]			
	DE: DE: Propranololiberin II precursor	[Contains: Gonadoliberin II (LHRH II) (Luteinizing hormone releasing hormone II) (Gonadotropin releasing hormone II) (GnRH II); GnRH-associated peptide III; GNRH2]			
	GN: Tupaia glis belangeri	(Common tree shrew).			
	OS: Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi; OC: Mammalia; Eutheria; Scandentia; Tupaiidae; Tupaias; NCBL_TaxID=9336;				
	OX: OX: NCBL_TaxID=9336;				
	RN: RN: SEQUENCE FROM N.A.				
	RC: RC: TISSUE=HYPOTHALAMUS; MEDLINE=97079039; PubMed=8921350;				
	RA: Kasten T.L., White S.A., Norton T.T., Bond C.T., Adelman J.P., Fernald R.D.,				
	RA: RT: Characterization of two new preprognRH mRNAs in the tree shrew: a first direct evidence for mesencephalic GnRH gene expression in a				
	RT: RT: placental mammal.", Gen. Comp. Endocrinol. 104:7-19(1996).				
	RL: CC: -!- FUNCTION: STIMULATES THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING HORMONES.				
	CC: CC: -!- SUBCELLULAR LOCATION: Secreted.				
	CC: CC: -!- TISSUE-SPECIFICITY: MIDBRAIN.				
	CC: CC: -!- SIMILARITY: BELONGS TO THE GNRH FAMILY.				
	CC: CC: This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to license@isb-sib.ch).				
	CC: CC: EMBL: U653327; AAB16838; 1; -.				
	DR: DR: InterPro: IPR00212; GnRH.				
	DR: DR: PROSITE: PS00446; GnRH; 1; PRam; PF00446; GnRH; 1;				
	KW: DR: Cleaveage on pair of basic residues; Hormone; Amidation; Signal.				
	FT: SIGNAL: FT: 25 BY SIMILARITY.				
	FT: CHAIN: FT: 26 PROGNADOLIBERIN II.				
	FT: PEPTIDE: FT: 35 GONADOLIBERIN II.				
	FT: PEPTIDE: FT: 39 GNRH-ASSOCIATED PEPTIDE II.				
	FT: MOD_RES: FT: 26 PYRROLIDONE CARBOXYLIC ACID				
	FT: MOD_RES: FT: 35 (BY SIMILARITY).				
	FT: SEQUENCE: SQ: 114 AA: 12123 MW: 680690E1C6869E1 CRC64;				
	Query Match Similarity: 8.9%; Score: 7; DB 1; Length: 114; Best Local Similarity: 100.0%; Pred. No. 2.8; Length: 114;				

Matches	7	PRPAAVP	13	0	Mismatches	0	Indels	0	Gaps	0	RESULT 3
OY											ZEST-DROME
Db	105	PRPAAVP	111	ID	ZEST-DROME	STANDARD;	PRT;	574 AA.			ID
				AC	P03956; 024996; 09V3FL;						AC
				P34474;							DT
				01-FEB-1994	(Rel. 28, created)						01-MAR-1989 (Rel. 10, Created)
				01-FEB-1995	(Rel. 31, Last sequence update)						01-NOV-1990 (Rel. 16, Last sequence update)
				DT	15-JUN-2002 (Rel. 41, Last annotation update)						16-OCT-2001 (Rel. 40, Last annotation update)
				DE	Hypothetical protein F58A4.7 in chromosome III.						DE Regulatory protein zeste.
				GN	F58A4.7.						Z OR EG-BAC19J11.3 OR CG7803.
				OS	Eukaryota: Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea; Caenorhabditis elegans.						OS Drosophila melanogaster (Fruit fly).
				OC	Rhabdiidae; Peledorinae; Caenorhabditis elegans.						OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephyoidea; Drosophilidae; Drosophila.
				OC	NCBI_TAXID=6239;						OC NCBI_TAXID=6227;
				RN	[1]						RN [1]
				RP	SEQUENCE FROM N.A.						RP SEQUENCE FROM N.A.
				RC	SPRAIN-BRISTOL N2;						RC MEDLINE=88112560; PubMed=3125410;
				RA	Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,						RA MEDLINE=94150718; PubMed=9706398;
				RA	Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,						RA MEDLINE=94150718; PubMed=9706398;
				RA	Craxton M., Dear S., Du Z., Durbin R., Favell A., Fraser A.,						RA MEDLINE=94150718; PubMed=9706398;
				RA	Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,						RA MEDLINE=94150718; PubMed=9706398;
				RA	Johnston L., Jones M., Kershaw J., Kirstein J., Laisteller N.,						RA MEDLINE=94150718; PubMed=9706398;
				RA	Latreille P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,						RA MEDLINE=94150718; PubMed=9706398;
				RA	Parsons J., Percy C., Rifkin L., Roopra A., Saunders D., Showkeen R.,						RA MEDLINE=94150718; PubMed=9706398;
				RA	Sims M., Smalldon N., Smith A., Smith M., Sonnhammer E., Staden R.,						RA MEDLINE=94150718; PubMed=9706398;
				RA	Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,						RA MEDLINE=94150718; PubMed=9706398;
				RA	Waterson N., Watson A., Weinstock L., Wilkinson-Sproat J.,						RA MEDLINE=94150718; PubMed=9706398;
				RA	Woldman P.,						RA MEDLINE=94150718; PubMed=9706398;
				RT	2.2 Mb of contiguous nucleotide sequence from chromosome III of c. elegans.";						RT MEDLINE=87218538; PubMed=3582372;
				RT	Nature 368:32-38(1994).						RT MEDLINE=87218538; PubMed=3582372;
				RT	- - - SUBCELLULAR LOCATION: Nuclear (Potential).						RT Pirrotta V., Manet E., Hardon E., Bickel S.E., Benson M.,
				RT	- - - SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF TRANSCRIPTION FACTORS. STRONGEST, TO TRANSCRIPTION FACTOR AP-4.						RT "Structure and sequence of the Drosophila zeste gene.";
				RT	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to license@isb-sib.ch).						RT EMBO J. 6:791-799(1987).
				RT	REVISIONS, AND SELF-ASSOCIATION						RT [3]
				RT	"Self-association of the Drosophila zeste protein is responsible for transvection effects";						RT REVISIONS, AND SELF-ASSOCIATION
				RT	EMBO J. 9:2959-2967(1990).						RT REVISIONS, AND SELF-ASSOCIATION
				RN	[4]						RT REVISIONS, AND SELF-ASSOCIATION
				RP	SEQUENCE FROM N.A.						RP SEQUENCE FROM N.A.
				RC	STRAIN-BERKELEY,						RC STRAIN-BERKELEY,
				RC	MEDLINE=20116006; PubMed=10731122;						RC MEDLINE=20116006; PubMed=10731122;
				RA	Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,						RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
				RA	Amanatides P.G., Scheer S.E., Li P.W., Roskins R.A., Galle R.F.,						RA Amanatides P.G., Scheer S.E., Li P.W., Roskins R.A., Galle R.F.,
				RA	George R.A., Lewis S.E., Richards S., Ashburner M., Henggeler S.N.,						RA George R.A., Lewis S.E., Richards S., Ashburner M., Henggeler S.N.,
				RA	Sutton G.G., Wortman J.R.R., Yandell M.D., Zhang Q., Chen L.X.,						RA Sutton G.G., Wortman J.R.R., Yandell M.D., Zhang Q., Chen L.X.,
				RA	Brandon R.C., Rogers Y.-H.-C., Blazquez R.G., Changue M., Pfeiffer B.D.,						RA Brandon R.C., Rogers Y.-H.-C., Blazquez R.G., Changue M., Pfeiffer B.D.,
				RA	Wan K.H., Doyle C., Baxter E.G., Heit G., Nelson C.R., Miklos G.L.G.,						RA Wan K.H., Doyle C., Baxter E.G., Heit G., Nelson C.R., Miklos G.L.G.,
				RA	Abril J.F., Agbayani A.A., An H.-J., Andrews-Pfankoch C., Baldwin D.,						RA Abril J.F., Agbayani A.A., An H.-J., Andrews-Pfankoch C., Baldwin D.,
				RA	Baillw R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,						RA Baillw R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
				RA	Besson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolstakow S.,						RA Besson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolstakow S.,
				RA	Borkova D., Botchan M.R., Bouck J., Brokschein P., Brottier P.,						RA Borkova D., Botchan M.R., Bouck J., Brokschein P., Brottier P.,
				RA	Burtik K.C., Busam D.A., Cadine E., Center A., Chandra I.,						RA Burtik K.C., Busam D.A., Cadine E., Center A., Chandra I.,
				RA	Cherry J.M., Cawley S., Danilke C., Davenport L.B., Davies P.,						RA Cherry J.M., Cawley S., Danilke C., Davenport L.B., Davies P.,
				RA	de Pablo B., Deicher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,						RA de Pablo B., Deicher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
				RA	Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,						RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
				RA	Durbin R.K.J., Evangelista C.C., Ferraz C., Ferreira J.A., Fleischmann W.,						RA Durbin R.K.J., Evangelista C.C., Ferraz C., Ferreira J.A., Fleischmann W.,
				RA	Frolier C., Gabrielian A.E., Gaqc N.S., Gebhart W.M., Glasser K.,						RA Frolier C., Gabrielian A.E., Gaqc N.S., Gebhart W.M., Glasser K.,
				RA	Gloedk A., Gong F., Correll J.H., Gu Z., Guan P., Harris M.,						RA Gloedk A., Gong F., Correll J.H., Gu Z., Guan P., Harris M.,
				RA	Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,						RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
				RA	Hostin N.L., Houston K.A., Howland T.J., Wei M.-H., Ihngwan C.,						RA Hostin N.L., Houston K.A., Howland T.J., Wei M.-H., Ihngwan C.,
				RA	Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,						RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
				RA	Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,						RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
				RA	Lesko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,						RA Lesko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
				RA	Liu X., Mattei B., McIntosh T.C.C., McLeod M.P., McPherson D.,						RA Liu X., Mattei B., McIntosh T.C.C., McLeod M.P., McPherson D.,
				RA	Markoulou G., Mishina N.V., Mobarry Morris J., Moskoff A.I.,						RA Markoulou G., Mishina N.V., Mobarry Morris J., Moskoff A.I.,
				RA	Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,						RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
				RA	Nelson D.R., Nelson K.A., Nixon J.M., Nusslein D.R., Facile J.,						RA Nelson D.R., Nelson K.A., Nixon J.M., Nusslein D.R., Facile J.,
				RA	Palazzo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,						RA Palazzo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
				RA	Rainert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,						RA Rainert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
				RA	Shue B.C., Siden-Kiamis I., Simpson M., Skupski M.P., Smith T.,						RA Shue B.C., Siden-Kiamis I., Simpson M., Skupski M.P., Smith T.,
				RA	Spler E., Spradling A.C., Stapleton M., Strong R., Sun E.,						RA Spler E., Spradling A.C., Stapleton M., Strong R., Sun E.,
				RA	Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,						RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
				RA	Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissenbach J.,						RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissenbach J.,
				RA	Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,						RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,

RA	Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,	FT DOMAIN 5 42	GLY-RICH.
RA	Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,	FT DOMAIN 152 430	GLN/ALA-RICH (OPA-REPEAT INVOLVED IN TRANSCRIPTIONAL ACTIVATION OR REPRESSION AT DIFFERENT TARGET LOCI) (POTENTIAL).
RA	Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;	FT	MISSING (IN STRAIN ME-K2).
RT	"the genome sequence of <i>Drosophila melanogaster</i> ;";	FT VARIANT 232 233	S -> A (IN REF. 4 AND 5).
RL	Science 287:2185-2195(2000).	FT CONFLICT 496 496	[5]
RP	SEQUENCE FROM N.A.	SQ SEQUENCE 574 AA;	61845 MW; 45DGD36C72F5CF3 CRC64;
RC	STRAIN=Oregon R;		
RX	MEDLINE=10196011; PubMed=10731137;		
RA	Benos P.V., Gatt M.K., Ashburner M., Murphy L., Harris D.,' Harris C.,' Demailles J., Cadieu E.,'	OY 9 PAAPVPP 15	Query Match 8.9%; Score 7; DB 1; Length 574;
RA	Barrel B.G., Ferraz C., Vidal S., Brun C.,' Mottier S., Galibert F.,' Borkova D.,'	RL 7; Conservative 0; Mismatches 0;	Best Local Similarity 100.0%; Pred. No. 11;
RA	Dreano S., Gloux S., Lejaire V., Mottier S., Galibert F.,' Borkova D.,'	DN 0; Indels 0; Gaps 0;	Matches 7;
RA	Minana B., Kafatos F.C., Louis C., Siden-Kiamos I., Bolshakov S.,' Papagiannakis G., Spanos L., Cox S., Madueno E., de Pablo B.,' Moudgil F.,' Modoell J., Peter A., Schoettler P., Werner M., Mourkioti F.,' Beinert N., Dowe G., Schaefer U., Jaekle H., Bucheton A.,' Callister D.M., Campbell L.A.,' Darlamusso A.,' Henderson N.S.,' McMillan P.J.,' Sailles C., Tait E.A.,' Valenti P.,' Saunders R.D.C.,' Glover D.M.;		
RT	"From sequence to chromosome: the tip of the X chromosome of D. melanogaster";		
RL	Science 287:2220-2222(2000).	[6]	
RP	SELF-ASSOCIATION.		
RX	MEDLINE=93259149; PubMed=8491197;		
RA	Chen J.D., Pirrotta V.;		
RT	*Multimerization of the <i>Drosophila zeste</i> protein is required for efficient DNA binding";		
RL	EMBO J. 12:2075-2083(1993).		
RN	[7]		
RP	SEQUENCE OF 56-323 FROM N.A.		
RC	STRAIN=ME-K1, ME-K2, ME-L1, ME-L12, ME-NJ1, and ME-NJ2;		
RX	MEDLINE=93360802; PubMed=8355601;		
RA	Hey J., Kilman R.M.;		
RT	*Population genetics and phylogenetics of DNA sequence variation at multiple loci within the <i>Drosophila melanogaster</i> species complex.>";		
RT	Mol. Biol. Evol. 10:804-822(1993).		
RL	-1- FUNCTION: INVOLVED IN TRANSECTION PHENOMENA (= SYNAPSIS-DEPENDENT GENE EXPRESSION), WHERE THE SYNAPTIC PAIRING OF CHROMOSOMES CARRYING GENES WITH WHICH ZESTE INTERACTS INFLUENCES THE EXPRESSION OF THESE GENES. ZESTE Binds TO DNA AND STIMULATES TRANSCRIPTION FROM A NEARBY PROMOTER.		
CC	MONOMERS.		
CC	-1- SUBCELLULAR LOCATION: Nuclear.		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to license@isb-sib.ch).		
CC	-1- SUBUNIT: SELF-ASSOCIATE FORMING COMPLEXES OF SEVERAL HUNDRED MONOMERS.		
CC	-1- SUBCELLULAR LOCATION: Nuclear.		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to license@isb-sib.ch).		
CC	TISSUE=Intestine;		
CC	MEDLINE=93266779; PubMed=10334869;		
CC	Prasad P.D., Wang H., Huang W., Fei Y.-J., Leibach F.H., Devoe L.D., Ganapathy V.;		
CC	"Molecular and functional characterization of the intestinal Na+-dependent multivitamin transporter";		
CC	Arch. Biochem. Biophys. 366:95-106(1999);		
CC	-1- FUNCTION: TRANSPORTS PANTOTHENATE, BIOTIN AND LIPOATE IN THE PRESENCE OF SODIUM.		
CC	-1- SUBCELLULAR LOCATION: Integral membrane Protein.		
CC	-1- SIMILARITY: BELONGS TO THE SODIUM:SOLUTE SYMPORTER FAMILY (SSF).		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to license@isb-sib.ch).		
CC	DR		
CC	EMBL; AF116241; AAD37502.1;		
CC	EMBL; AF059307; AAD31727.1;		
CC	EMBL; AF081571; AAD37481.1;		
CC	Genew; HGNC:11041; SLC5A6.		
CC	MIM: 604204;		
CC	InterPro: IPR001734; Na_solut_symport.		
DR	Pfam: PF00474; SSF_1.		
DR	TIGRFAMS; TIGR00813; sss_1.		
DR	EMBL; AF081571; AAD37481.1;		
DR	EMBL; AF081571; AAD37481.1;		
DR	Genew; HGNC:11041; SLC5A6.		
FT	TRANSFAC; T00918;		
DR	FlyBase; FBgn0004050; Z.		
DR	DNA-binding; transcription regulation; Nuclear protein; Polymorphism.		
FT	INIT_MET 0		
FT	DOMAIN 1 46		
FT	DNA_BIND 47 127		
	HYPROPHOBIC.		
	SPECIFIC, WITH ZESTE LOCUS.		

KW Transport; transmembrane; Sodium transport; Symport; Glycoprotein.  
 FT TRANSMEM 24 44 POTENTIAL.  
 FT TRANSMEM 68 88 POTENTIAL.  
 FT TRANSMEM 101 121 POTENTIAL.  
 FT TRANSMEM 143 163 POTENTIAL.  
 FT TRANSMEM 176 196 POTENTIAL.  
 FT TRANSMEM 199 219 POTENTIAL.  
 FT TRANSMEM 256 276 POTENTIAL.  
 FT TRANSMEM 297 317 POTENTIAL.  
 FT TRANSMEM 336 356 POTENTIAL.  
 FT TRANSMEM 396 416 POTENTIAL.  
 FT TRANSMEM 428 448 POTENTIAL.  
 FT TRANSMEM 456 476 POTENTIAL.  
 FT TRANSMEM 528 548 POTENTIAL.  
 FT CARBOHYD 138 138 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 489 489 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT SEQUENCE 635 AA; 68701 MW; 4F10369A3916564F CRC64;  
 SQ 111111 9 APSPSTS 15  
 Db 46 APSPSTS 52  
 Qy |||||||

Query Match 8.9%; Score 7; DB 1; Length 635;  
 Best Local Similarity 100.0%; Pred. No. 12;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 5  
 GC03\_MOUSE STANDARD; PRT; 677 AA.  
 AC P8854;  
 DT 15-JUN-2002 (Rel. 41, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Gamma-tubulin complex component 3 (GCP-3) (Fragment).  
 GN GC03.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Cetartiodactyla; Homo.  
 OX NCBI\_TaxID=6666;  
 RN [1] SEQUENCE FROM N.A. (ISOFORM 1).  
 RN TISSUE=cervical carcinoma;  
 RX MEDLINE=98234402; PubMed=9566967;  
 RA Murphy S.M., Urbani L., Stearns T.;  
 RT "The mammalian gamma-tubulin complex contains homologues of the yeast  
 RT spindle pole body components spc97p and spc98p";  
 RL J. Cell Biol. 141:663-674(1998).  
 RN [2] SEQUENCE FROM N.A. (ISOFORMS 1 AND 2), AND CHARACTERIZATION.  
 RX MEDLINE=98234404; PubMed=9566969;  
 RA Taissin A.-M., Celati C., Moudjou M., Bornens M.;  
 RT Characterization of the human homologue of the yeast spc98p and its  
 RT association with gamma-tubulin.";  
 RL J. Cell Biol. 141:689-701(1998).  
 RN [3] SEQUENCE FROM N.A. (ISOFORMS 2 AND 3).  
 RP TISSUE=Eye, and Muscle;  
 RA Strausberg R.;  
 CC Submitted (SIEP-2001) to the EMBL/GenBank/DBJ databases.  
 -!- FUNCTION: Gamma-tubulin complex is necessary for microtubule  
 nucleation at the centrosome.  
 CC -!- SUBUNIT: Gamma-tubulin complex is composed of gamma-tubulin, GCP2,  
 CC -!- SUBCELLULAR LOCATION: Centrosome (By similarity).  
 CC -!- SIMILARITY: BELONGS TO THE GCP FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 or send an email to license@isb-sib.ch).  
 CC ---  
 DR EMBL; AR042378; AAC39127.1; -  
 DR EMBL; AT003061; CAA05932.1; -  
 DR EMBL; AU003002; CAA05833.1; -  
 DR EMBL; BC013781; AAH13781.1; -  
 DR EMBL; BC007763; AAH07763.1; -  
 KW Microtubules; Alternative splicing.  
 FT DOMAIN 825 829 POLY-GLU.  
 FT VARSPLIC 817 824 GQGWTAA > VEMCLXCV (IN ISOFORM 2).  
 FT VARSPLIC 825 907 MISSING (IN ISOFORM 2).  
 FT VARSPLIC 391 434 RKGELASRAVHATGPyMPSIWLWHTLSIVSHPVISFLY  
 FT RWT > PTVFVTPWFTTRDPTRDFMMVPPTRVFFPTV  
 FT WHSLCFFRL (IN ISOFORM 3).  
 FT VARSPLIC 435 907 MSING (IN ISOFORM 3).  
 FT CONFLICT 208 208 T -> S (IN REF. 2).  
 FT CONFLICT 361 361 S -> I (IN REF. 2).  
 SQ SEQUENCE 907 AA; 103570 MW; 70FE2FD8C80344D CRC64;  
 FT VARSPLIC 435 907  
 FT CONFLICT 208 208  
 FT CONFLICT 361 361  
 SQ SEQUENCE 907 AA; 103570 MW; 70FE2FD8C80344D CRC64;  
 Query Match 8.9%; Score 7; DB 1; Length 907;  
 Best Local Similarity 100.0%; Pred. No. 13;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 194 HPVLISFL 200

Best Local Similarity 100.0%; Pred. No. 17; Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 7  
 PSBK\_PINTH STANDARD; PRT; 56 AA.  
 ID PSBK\_PINTH  
 AC P41598;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE photosystem II reaction center protein K precursor (PSII-K).  
 GN PSBK  
 OS Pinus thunbergii (Green pine) (Japanese black pine).  
 OG Chloroplast.  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Conilleropsida; Coniferales; Pinaceae; Pinus.  
 NCBI\_TAXID=3350;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. BELO; TISSUE=Tassel;  
 RX MEDLINE=94004987; PUBMED=8401606;  
 RA Wright S.Y., Suner M.-M., Bell P.J., Vaudin M., Greenland A.J.;  
 RT "Isolation and characterization of male flower cDNAs from maize";  
 RL Plant J. 3:41-49(1993).  
 CC -1- TISSUE SPECIFICITY: ENHANCED EXPRESSION IN MALE FLOWERS.  
 CC -1- ACCUMULATES IN THE TAPETUM.  
 CC -1- DEVELOPMENTAL STAGE: ASSOCIATED WITH MICROSPOROGENESIS AND DECLINES AS MATURE POLLEN IS PRODUCED.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC CC  
 CC DR EMBL: X67323; CAA47737.1; ~.  
 DR PIR: S25104; S25104.  
 DR Maizedb: 69183; ~.  
 DR Signal.  
 FT SIGNAL 1 23 OR 24, OR 25 (POTENTIAL).  
 RA Sugiura M.; Tsudzuki J., Ito S., Nakashima K., Tsudzuki T.,  
 RT "Loss of all ndh genes as determined by sequencing the entire chloroplast genome of the black pine *Pinus thunbergii*";  
 RL Proc. Natl. Acad. Sci. U.S.A. 91:9794-9798(1994).  
 CC -1- FUNCTION: THIS PROTEIN IS A COMPONENT OF THE REACTION CENTER OF PHOTOSYSTEM II.  
 CC -1- SIMILARITY: BELONGS TO THE PSBK FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC CC  
 CC DR EMBL: D17510; BAA04312.1; ALT INIT.  
 DR InterPro: IPRO03687; PSII\_PSBK.  
 DR Pfam: PF02533; Psbk; 1.  
 KW Photosystem II; Chloroplast.  
 FT PROPEP 1 19 POTENTIAL.  
 FT CHAIN 20 56 PHOTOSYSTEM II REACTION CENTER PROTEIN K.  
 SQ SEQUENCE 56 AA; 6346 MW; 18B1DBEFL198ACASA CRC64;  
 Query Match 7.6%; Score 6; DB 1; Length 56;  
 Best Local Similarity 100.0%; Pred. No. 16;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 PVLSPFL 30  
 DB 39 PVLSPFL 44

RESULT 8  
 MF14\_MAIZE STANDARD; PRT; 126 AA.  
 ID MF14\_MAIZE  
 AC Q01900;  
 DT 01-JUL-1993 (Rel. 26, Created)  
 DT 01-JUL-1993 (Rel. 26, Last sequence update)  
 DT 01-JUN-1994 (Rel. 29, Last annotation update)  
 DE MF14 protein precursor.  
 GN MF14.  
 Zea mays (Maize); Eukaryota; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; Zea; NCBI\_TAXID=4577;

RESULT 9  
 VP3\_BPCHP STANDARD; PRT; 145 AA.  
 ID VP3\_BPCHP  
 AC P19194;  
 DT 01-NOV-1990 (Rel. 16, Created)  
 DT 01-NOV-1990 (Rel. 16, Last sequence update)  
 DT 01-NOV-1990 (Rel. 16, Last annotation update)  
 DE Protein VP3 (ORF3).  
 OS Bacteriophage Chp1.  
 OC Viruses; ssDNA viruses; Microviridae; Microvirus.  
 OX NCBI\_TAXID=12267;  
 RN [1]  
 RP SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.  
 RX MEDLINE=9011116; PUBMED=2607341;  
 RA Storey C.C., Lusher M.M., Richmond S.J.;  
 RT "Analysis of the complete nucleotide sequence of Chp1, a phage which infects avian Chlamydia psittaci";  
 RL J. Gen. Virol. 70:3381-3390(1989).  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC CC  
 CC DR EMBL: D00624; BAA00509.1; ~.  
 DR PIR: JU0347; JU0347.  
 DR Signal.  
 KW SEQUENCE 145 AA; 16601 MW; 6B4DFA64663587CD CRC64;  
 SQ Sequence 145 AA; 16601 MW; 6B4DFA64663587CD CRC64;  
 Query Match 7.6%; Score 6; DB 1; Length 145;  
 Best Local Similarity 100.0%; Pred. No. 36;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 39 AYVSLP 44  
 ID |||||  
 AC 001636;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Laminin beta-1 chain variant (laminin beta-1-2 chain) (Fragment).  
 OS Gallus gallus (Chicken).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
 OC Gallus.  
 OX NCBI\_TaxID=9031;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE="eye";  
 RX MEDLINE=93015947; PubMed=1400373;  
 RA O'Rear J.J.;  
 RT "A novel laminin B1 chain variant in avian eye.";  
 RL J. Biol. Chem. 267:20555-20557(1992).  
 CC -1- FUNCTION: Binding to cells via a high affinity receptor, laminin  
 CC is thought to mediate the attachment, migration, and organization  
 CC of cells into tissues during embryonic development by interacting  
 CC with other extracellular matrix components.  
 CC -1- SUBUNIT: Laminin is a complex glycoprotein, consisting of three  
 CC different polypeptide chains (alpha, beta, gamma), which are bound  
 CC to each other by disulfide bonds into a cross-shaped molecule  
 CC Comprising one long and three short arms with globules at each  
 CC end  
 CC -1- SUBCELLULAR LOCATION: Extracellular.  
 CC -1- COMPONENT): CONTAINS 1 LAMININ N-TERMINAL DOMAIN (MAJOR  
 CC -1- SIMILARITY: BELONGS TO THE POLY/OUTJ/APSJ/EXEJ/XCPW FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to license@isb-sib.ch).  
 DR EMBL; AA000409; AAC76356.1; -.  
 DR EcGene; EG1289; 9spJ.  
 DR InterPro; IPR001120; Prok\_N\_methyltn.  
 DR PROSITE; PS00409; PROK\_AR\_NTER\_METHYL; 1.  
 KW Transport; Methylation; Complete proteome.  
 FT PROPEP 1 7 BY SIMILARITY.  
 FT CHAIN 8 195 PROBABLE GENERAL SECRETION PATHWAY  
 FT PROTEIN J. METHYLATION (BY SIMILARITY).  
 FT MOD\_RES 8 8 195 AA; 22231 MW; 4A25DB082106B1C5 CRC64;  
 SQ SEQUENCE 198 AA; 21830 MW; 6FD669761B92C442 CRC64;  
 Query Match 7.6%; Score 6; DB 1; Length 198;  
 Best Local Similarity 100.0%; Pred. No. 47;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 44 PLAPLS 49  
 Db 92 PLAPLS 97  
 RESULT 11  
 LMHV\_CHICK STANDARD; PRT; 198 AA.  
 ID LMHV\_CHICK  
 AC 001636;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Laminin beta-1 chain variant (laminin beta-1-2 chain) (Fragment).  
 OS Gallus gallus (Chicken).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
 OC Gallus.  
 OX NCBI\_TaxID=9031;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=95141096; PubMed=8555498;  
 RA Ma C.; Staudt L.M.;  
 RT "LAF-4 encodes a lymphoid nuclear protein with transactivation  
 potential that is homologous to AF-4, the gene fused to MLL in  
 t(4;11) leukemias.";



This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC prophage

CC

RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;

RN "Complete genome sequence of enterohemorrhagic Escherichia coli

RN 0157:H7 and genomic comparison with a laboratory strain K-12.";

RN DNA Res. 8:11-22(2001).  
[4]

RN SEQUENCE OF 98-209 FROM N.A.

RN STRAIN=K12;

RN Imanura R., Niki H., Yamamoto K., Ogura T., Fujita N., Ishihama A.,

RN Hirata S.;

RN Submitted (JUN-1993) to the EMBL/genBank/DDJB databases.

RN [5]

RN CHARACTERIZATION.

RN STRAIN=K12 / MG1655;

RX MEDLINE-20011437; PubMed-10542272;

RX Dunn C.A., O'Handley S.F., Frick D.N., Bessman M.J.;

RX "Studies on the ADP-ribose pyrophosphatase subfamily of the nudix hydrolases and tentative identification of trgb, a gene associated with tellurite resistance.";

RX J. Biol. Chem. 274:32318-32324(1999).

RN -I- CATALYTIC ACTIVITY: ADP-ribose + H<sub>2</sub>O = AMP + D-ribose 5-

RN -- CORACOR: REQUIRES DIVALENT IONS: MAGNESIUM, MANGANESE OR ZINC.

RN -I- SIMILARITY: BELONGS TO THE NUDIX HYDROLASE FAMILY. NUDIX SUBFAMILY.

RN CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

RN CC

RN DR EMBL: U28377; AAA620202; 1;

RN DR EMBL: AE00385; AAC676070; 1;

RN DR EMBL: AE005533; AAC58173; 1;

RN DR EMBL: AP002564; BAB37345; 1;

RN DR EMBL: DI6557; ; NOT\_ANNOTATED\_CDS.

RN DR Ecogene; ECG12184; nudF.

RN DR InterPro; IPRO04385; Cons\_hypoth52.

RN DR InterPro; IP000085; NUDIX\_hydrolase.

RN DR PRAM; PF00293; NUDIX; 1.

RN DR PRINTS; PR00502; NUDIX\_FAMILY.

RN DR TIGR3MS; TIGR00052; Cons\_hypoth52; 1.

RN DR PROSITE; PS0083; NUDIX; 1.

RN DR Hydrolase; MAGNESIUM; manganese; Zinc; Cadmium

FT DOMAIN 97/ 118 NUDIX BOX.

SQ SEQUENCE 209 AA; 23667 MW; 2CF7E9d63B9615 CRC64;

Query Match 7.6%; Score 6; DB 1; Length 209;

Best Local Similarity 100.0%; Pred. No. 50;

Matches 6; Conservative 0; Mismatches 0;

Indels 0; Gaps 0;

Db 126 PVLSTL 131

QY 25 PVLSTL 30

QY 11111

QY 11111

Db 126 PVLSTL 131

RN SEQUENCE FROM N.A.

RC STRAIN=0157:H7 / EDL933 / ATCC 700927;

RX MEDLINE-21074935; PubMed-11206551;

RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D., Riley M., Collado-Vides J., Glazner J.D., Rose C.K., Mayhew G.F., Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J., Mau B., Shao Y., Hackett J., Klink S., Boutin A., Shao Y., Miller L., Grotbeck E.J., Davis N.W., Lim A., Dimanlana E.T., Pottamousis K., Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C., Welch R.A., Blattner F.R.;

RT "The complete genome sequence of Escherichia coli K-12.";

RT Science 277:1453-1474(1997).

RL RN [2]

RC SEQUENCE FROM N.A.

RC STRAIN=0157:H7 / RIMD 0509952;

RX MEDLINE-21156231; PubMed-11258796;

RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K., Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T., Iida T., Takami H., Honda T., Sakakawa C., Ogasawara N., Yasunaga T.,

RA "Genome sequence of enterohaemorrhagic Escherichia coli 0157:H7.";

RL RN [3]

RC SEQUENCE FROM N.A.

RC STRAIN=0157:H7 / RIMD 0509952;

RX MEDLINE-21156231; PubMed-11258796;

RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K., Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T., Iida T., Takami H., Honda T., Sakakawa C., Ogasawara N., Yasunaga T.,

RA "Genome sequence of enterohaemorrhagic Escherichia coli 0157:H7.";

RL RN [3]

Search completed: January 14, 2003, 17:13:20  
Job time : 5.69679 secs

Copyright (c) 1993 - 2003 Compugen Ltd.

Om protein - protein search, using sw model

Run on: January 14, 2003, 17:11:37 ; search time 4.6004 seconds  
 (without alignments)  
 1650.860 Million cell updates/sec

Title: US-09-506-079E-1

Perfect score: 79

Sequence: 1 GAXHXXPRPAAVPVPXIQP.....VGRGXDPDAHVAVLRSREG 79

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 283224 seqs, 96134422 residues

Word size : 0

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : PIR7.3:  
 1: pir1:  
 2: pir2:  
 3: pir3:  
 4: pir4:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query	Match Length	DB ID	Description
1	7	8.9	121	2 AH1848	hypothetical protein
2	7	8.9	292	2 SA0979	hypothetical protein
3	7	8.9	382	2 AB8561	protein F58A4.7b
4	7	8.9	389	2 AG7577	hypothetical protein
5	7	8.9	389	2 AH2797	conserved hypothetical protein
6	7	8.9	436	2 B70473	protochlorophyllide oxygenase
7	7	8.9	495	2 TI17478	hypothetical protein
8	7	8.9	520	2 CT0311	hypothetical protein
9	7	8.9	555	2 A26639	regulatory protein
10	7	8.9	595	2 SU1272	hypothetical protein
11	7	8.9	695	2 TI15862	hypothetical protein
12	7	8.9	865	2 AB5032	hypothetical protein
13	7	8.9	879	2 SA9910	chloroplast outer membrane protein
14	7	8.9	1503	2 T01098	chloroplast outer membrane protein
15	7	8.9	2240	2 T37057	probable multi-domain protein
16	6	7.6	59	2 T07432	photosystem II protein
17	6	7.6	65	2 CT2558	hypothetical protein
18	6	7.6	74	2 FR2764	hypothetical protein
19	6	7.6	98	2 S68136	NADH2 dehydrogenases
20	6	7.6	120	2 DB31166	hypothetical protein
21	6	7.6	122	2 AB2199	hypothetical protein
22	6	7.6	126	2 S25104	MSF14 protein - maize
23	6	7.6	128	2 GI1264	conserved hypothetical protein
24	6	7.6	128	2 AB7510	hypothetical protein
25	6	7.6	129	2 BB3173	hypothetical protein
26	6	7.6	142	2 C38723	hypothetical protein
27	6	7.6	142	2 BT2571	hypothetical protein
28	6	7.6	145	1 JU0347	capsid protein vp3
29	6	7.6	163	2 AB2025	hypothetical protein

30 6 7.6 171 2 TA3959 hypothetical protein  
 31 6 7.6 182 2 S76346 hypothetical protein  
 32 6 7.6 182 2 FR9221 hypothetical protein  
 33 6 7.6 191 2 T04853 probable general s laminin B1 chain v unknown protein en hypothetical protein  
 34 6 7.6 195 2 F65126 capsular polysaccharide ADP-ribose diphosphorylase diphosphorylase  
 35 6 7.6 198 2 A55067 ADP-ribose diphosphorylase diphosphorylase  
 36 6 7.6 205 2 FR5515 ADP-ribose diphosphorylase diphosphorylase  
 37 6 7.6 205 2 C90665 cytochrome-c oxidase  
 38 6 7.6 206 2 E43335 capsular polysaccharide  
 39 6 7.6 207 2 A69941 ADP-ribose diphosphorylase diphosphorylase  
 40 6 7.6 209 2 A65964 ADP-ribose diphosphorylase diphosphorylase  
 41 6 7.6 209 2 B91119 ADP-ribose diphosphorylase diphosphorylase  
 42 6 7.6 209 2 H65090 probable transcript  
 43 6 7.6 210 2 T21742 probable transcript  
 44 6 7.6 214 2 T22087 probable transcript  
 45 6 7.6 215 2 S78307 probable transcript

## ALIGNMENTS

RESULT 1

AH1848 hypothetical protein all0337 [imported] - Nostoc sp. (strain PCC 7120)

C;Species: Nostoc sp.  
 C;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120  
 C;Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 30-Jun-2002  
 C;Accession: AH1848  
 R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanaabe, A.; Irigu Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takarawa, M.; Yamada, M.; Yasuda, M.; Tabata DNA Res. 8, 205-213, 2001  
 A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Nostoc sp. str. PCC 7120  
 A;Reference number: AB1807; MUID:21595285; PMID:11759840  
 A;Accession: AH1848  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-121 <KUR>  
 A;Cross-references: GB:BA000019; PIDN:BAB72295.1; PID:g17129682; GSPDB:GN00179  
 A;Experimental source: strain PCC 7120  
 C;Genetics:  
 C;Gene: all0337

RESULT 2

Query Match Score 8.9%; DB 2; Length 121; Best Local Similarity 10.0%; Pred. No. 5.3; Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 25 PVLFLR 31

Db 56 PVLFLR 62

S40979 hypothetical protein F58A4.7 - Caenorhabditis elegans  
 C;Species: Caenorhabditis elegans  
 C;Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 23-Feb-1997  
 C;Accession: S40979  
 R;Berks, M.  
 submitted to the EMBL Data Library, February 1997  
 A;Reference number: S40973  
 A;Accession: S40979  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-292 <BER>  
 A;Cross-references: EMBL:Z22179  
 C;Genetics:  
 A;Introns: 60/2; 133/1; 273/1  
 C;Keywords: nucleus

Query Match Score 8.9%; DB 2; Length 292; Best Local Similarity 10.0%; Pred. No. 12; Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 46 APSPSTS 52 ;  
 QY 48 APSPSTS 54 ;  
 Db ;  
 C;Species: *Caenorhabditis elegans*  
 C;Accession: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 10-May-2001  
 R;anonymous' The *C. elegans* sequencing Consortium.  
 Science 282, 2012-2018, 1998  
 A;Reference sequence of the nematode *C. elegans*: a platform for investigating biolog  
 A;Reference number: A75000; MURB:9909613; PMID:9851916  
 A;Note: see websites genome.wustl.edu/gsc/C\_elegans/ and www.sanger.ac.uk/Projects/C\_eleg  
 A;Accession: B88561  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-382 <STO>  
 A;Cross-references: GB:chr\_-III; PIDN:CAA80170.1; PID:93877829; GSPDB:GN00021; CESP:F58A4  
 C;Genetics:  
 A;Gene: F58A4\_7b  
 A;Map position: 3  
 Query Match 8.9%; Score 7; DB 2; Length 382;  
 Best Local Similarity 100.0%; Pred. No. 15;  
 Matches 7; Conservative 0; Mismatches 0;  
 Indels 0; Gaps 0;  
 QY 46 APSPSTS 52 ;  
 Db 48 APSPSTS 54 ;  
 RESULT 4  
 A93577 hypothetical protein AGR\_C\_3311 [imported] - Agrobacterium tumefaciens (strain C58, Cere  
 C;Species: Agrobacterium tumefaciens  
 C;Date: 30-Sep-2001 #sequence\_revision 30-Sep-2001 #text\_change 11-Jan-2002  
 C;Accession: A93577  
 R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Ourolo, B.; Goldman, A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.; Science 294, 2323-2328, 2001  
 A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum  
 A;Reference number: A97359; PMID:11743194  
 A;Accession: A97577  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-389 <KUR>  
 A;Cross-references: GB:AE007869; PIDN:AAK87570.1; PID:gi15156910; GSPDB:GN00169  
 C;Genetics:  
 A;Gene: AGR\_C\_3311  
 A;Map position: circular chromosome  
 Query Match 8.9%; Score 7; DB 2; Length 389;  
 Best Local Similarity 100.0%; Pred. No. 15;  
 Matches 7; Conservative 0; Mismatches 0;  
 Indels 0; Gaps 0;  
 QY 9 PANVPVP 15 ;  
 Db 160 PANVPVP 166 ;  
 RESULT 5  
 AH2797 conserved hypothetical protein Atul800 [imported] - Agrobacterium tumefaciens (strain C5  
 C;Species: Agrobacterium tumefaciens  
 C;Accession: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 11-Jan-2002  
 R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, T.; erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kulyavin, T.; Levy, R.; Li, M.; McClell ; Karp, P.; Romero, P.; Zhang, S.  
 A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kam ster, E.W.  
 A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.  
 A;Reference number: AB2577; PMID:11743193  
 A;Accession: AH2797  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-389 <KUR>  
 A;Cross-references: GB:AE008688; PIDN:AAL42798.1; PID:gi17740243; GSPDB:GN00186  
 A;Experimental source: strain C58 (Dupont)  
 C;Genetics:  
 A;Gene: Atul800  
 A;Map position: circular chromosome  
 Query Match 8.9%; Score 7; DB 2; Length 389;  
 Best Local Similarity 100.0%; Pred. No. 15;  
 Matches 7; Conservative 0; Mismatches 0;  
 Indels 0; Gaps 0;  
 QY 9 PAAVPVP 15 ;  
 Db 160 PAAVPVP 166 ;  
 RESULT 6  
 B70473 hypothetical protein oxidase - *Aquifex aeolicus*  
 C;Species: *Aquifex aeolicus*  
 C;Date: 08-May-1998 #sequence\_revision 08-May-1998 #text\_change 05-Nov-1999  
 C;Accession: B70473  
 R;Beckert, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; V. Nature 392, 353-358, 1998  
 A;Title: The complete genome of the hyperthermophilic bacterium *Aquifex aeolicus*.  
 A;Reference number: A70300; MURB:98196666; PMID:9537320  
 A;Accession: B70473  
 A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A;Molecule type: DNA  
 A;Residues: 1-436 <AQF>  
 A;Cross-references: GB:AE000768; NID:92984249; PIDN: AAC0777.1; PID:gi2984251; GB:AE00  
 A;Experimental source: strain VF5  
 C;Genetics:  
 A;Gene: hemG  
 Query Match 8.9%; Score 7; DB 2; Length 436;  
 Best Local Similarity 100.0%; Pred. No. 17;  
 Matches 7; Conservative 0; Mismatches 0;  
 Indels 0; Gaps 0;  
 QY 26 VLSFLRP 32 ;  
 Db 428 VLSFLRP 434 ;  
 RESULT 7  
 TI1478 hypothetical protein PCZA361\_11 - *Amycolatopsis orientalis*  
 C;Species: *Amycolatopsis orientalis*  
 C;Date: 02-Sep-2000 #sequence\_revision 02-Sep-2000 #text\_change 02-Sep-2000  
 C;Accession: TI1478  
 R;Van Wageningen, A.; Kirkpatrick, P.; Williams, D.; Harris, B.; Kershaw, J.; Lennard Chem. Biol. 3, 155-162, 1998  
 A;Title: Sequencing and analysis of genes involved in the biosynthesis of a vancomyci  
 A;Reference number: 218804  
 A;Accession: TI1478  
 A;Status: preliminary; translated from GB/EMBL/DDBJ  
 A;Molecule type: DNA  
 A;Residues: 1-495 <VAN>  
 A;Cross-references: EMBL:AU223998; NID:e1251208; PID:e1251217; PIDN:CAA11769.1  
 Query Match 8.9%; Score 7; DB 2; Length 495;  
 Best Local Similarity 100.0%; Pred. No. 19;  
 Matches 7; Conservative 0; Mismatches 0;  
 Indels 0; Gaps 0;

Qy 7 PRPAAVP 13  
 Qy |||||  
 Db 27 PRPAAVP 33

**RESULT 8**

C70311 hypothetical protein aq\_116 - Aquifex aeolicus  
 C;Species: Aquifex aeolicus  
 C;Accession: 08-May-1998 #sequence\_revision 08-May-1998 #text\_change 09-Jun-2000  
 R;Deckert, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; Ov  
 Nature 392, 353-358, 1998  
 A;Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.  
 A;Accession: C70311  
 A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A;Molecule type: DNA  
 A;Residues: 1-520 <XOF>  
 A;Experimental source: strain VF5  
 A;Genetics:  
 C;Superfamily: Aquifex aeolicus hypothetical protein aq\_116  
 Query Match 8.9%; Score 7; DB 2; Length 520;  
 Best Local Similarity 100.0%; Pred. No. 20;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 47 PLSPTSV 53  
 Db 2 PLSPTSV 8

**RESULT 9**

A26639 regulatory protein zeste - fruit fly (Drosophila sp.)  
 C;Species: Drosophila sp.  
 C;Date: 16-Aug-1988 #sequence\_revision 16-Aug-1988 #text\_change 16-Feb-1997  
 C;Accession: A26639  
 R;Pirootta, V.; Manet, E.; Hardon, E.; Bickel, S.E.; Benson, M.  
 EMBL J. 6, 791-799, 1987  
 A;Title: Structure and sequence of the Drosophila zeste gene.  
 A;Reference number: A26639; MUID:87218538; PMID:3582372  
 A;Accession: A26639  
 A;Molecule type: DNA  
 A;Residues: 1-555 <PIR>  
 C;Genetics:  
 A;Gene: zeste  
 A;Cross-references: FlyBase:FBgn0004050  
 A;Introns: 103/2; 284/1  
 C;Keywords: DNA binding; transcription regulation

Query Match 8.9%; Score 7; DB 2; Length 555;  
 Best Local Similarity 100.0%; Pred. No. 21;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 9 PAAVVP 15  
 Db 430 PAAVVP 436

**RESULT 10**

S01272 regulatory protein zeste - fruit fly (Drosophila melanogaster)  
 C;Species: Drosophila melanogaster  
 C;Date: 30-Jun-1989 #sequence\_revision 30-Jun-1989 #text\_change 21-Jul-2000  
 C;Accession: S01272  
 R;Mansukhani, A.; Gunaratne, P.H.; Sherwood, P.W.; Sneath, B.J.; Goldberg, M.L.  
 Mol. Gen. Genet. 211, 121-128, 1988  
 A;Title: Nucleotide sequence and structural analysis of the zeste locus of Drosophila me  
 A;Reference number: S01272; MUID:88142560; PMID:3125410

Qy 7 PRPAAVP 13  
 Qy |||||  
 Db 27 PRPAAVP 33

**RESULT 11**

TJ5862 hypothetical protein C56E6.5 - Caenorhabditis elegans  
 C;Species: Caenorhabditis elegans  
 C;Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 20-Sep-1999  
 C;Accession: TJ5862  
 R;Fulton, L.  
 submitted to the EMBL data library, November 1995  
 A;Description: The sequence of C. elegans cosmid C56E6.  
 A;Accession number: S69019  
 A;Status: preliminary; translated from GB/EMBL/DDBJ  
 A;Molecule type: DNA  
 A;Residues: 1-595 <FUL>  
 A;Cross-references: EMBL:U39996; NID:gi1055114; PID:gi1055118; PIDN:AAA81092.1; CESP:C5  
 C;Genetics:  
 A;Gene: CESP:C56E6.5  
 A;Introns: 23/2; 88/1; 134/3; 174/1; 265/1; 293/3; 362/2; 392/1; 436/2; 494/2; 539/2  
 Query Match 8.9%; Score 7; DB 2; Length 595;  
 Best Local Similarity 100.0%; Pred. No. 23;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 26 VLSELRLP 32  
 Db 357 VLSELRLP 363

**RESULT 12**

A85032 hypothetical protein At4g02510 [imported] - Arabidopsis thaliana  
 C;Species: Arabidopsis thaliana (mouse-ear cress)  
 C;Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 16-Feb-2001  
 C;Accession: A85032  
 R;anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Sp  
 Nature 402, 769-771, 1999  
 A;Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.  
 A;Reference number: A85001; MUID:20083488; PMID:10617198  
 A;Accession: A85032  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-665 <STO>  
 A;Cross-references: GB:NC\_001268; NID:gi7269011; PIDN:CA80744.1; GSPDB:GN00140  
 C;Genetics:  
 A;Gene: At4g02510  
 A;Map position: 4

Query Match 8.9%; Score 7; DB 2; Length 665;  
 Best Local Similarity 100.0%; Pred. No. 32;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 9 PAAVVP 15  
 Qy |||||

Db 584 PAAVPVP 590

RESULT 13

S49910 chloroplast outer envelope protein OEP86 precursor - garden pea

C;Species: *Psium sativum* (garden Pea)

C;Date: 26-Jan-1995 #sequence\_revision 10-Feb-1995 #text\_change 02-Feb-2001

R;Hirsch, S.; Soll, J.

submitted to the EMBL Data Library, March 1994

A;Reference number: S49910

A;Accession: S49910

A;Molecule type: mRNA

A;Residues: I-L-879 &lt;HIR&gt;

A;Cross-references: EMBL:Z31581; NID:9599957; PID:9599958

R;Hirsch, S.; Muckel, B.; Heemeier, F.; von Heijne, G.; Soll, J.

Science 266, 1989-1992, 1994

A;Title: A receptor component of the chloroplast protein translocation machinery.

A;Reference number: A55386; MUID:9509324; PMID:7801125

A;Status: nucleic acid sequence not shown

A;Molecule type: mRNA

A;Residues: I-137, Q-139-875, 877-879 &lt;HIR&gt;

R;Kessler, F.; Blobel, G.; Patel, H.A.; Schnell, D.J.

Science 266, 1035-1039, 1994

A;Title: Identification of two GTP-binding proteins in the chloroplast protein import ma

A;Reference number: A55171; MUID:95063938; PMID:7973656

A;Accession: A55171

A;Molecule type: mRNA

A;Residues: I-372, EQQ-, 376-879 &lt;KES&gt;

A;Cross-references: GB:L36557; NID:9576508; PIDN:AA153276.1; PID:9576509

C;Keywords: chloroplast; GTP binding; membrane protein; nucleotide binding; P-loop

F;24-252/Region: nucleotide-binding motif A (P-loop)

Query Match 8.9%; Score 7; DB 2; Length 1503;  
Best Local Similarity 100.0%; Pred. No. 53;  
Matches 7; Conservative 0; Mismatches 0;  
Indels 0; Gaps 0;

Qy 9 PAAVPVP 15

Db 599 PAAVPVP 605

RESULT 14

T01098 chloroplast outer envelope protein OSP86 homolog T10P11.19 - *Arabidopsis thaliana*C;Species: *Arabidopsis thaliana* (mouse-ear cress)

C;Date: 12-Feb-1999 #sequence\_revision 12-Feb-1999 #text\_change 02-Feb-2001

C;Accession: T01098; T01299

R;Kaplan, N.; Johnson, D.; Schutz, K.; Gnoj, L.; Hoffman, J.; Till, S.; de la Bastide, M.

hi, M.; Martenssen, R.; Chen, B.Y.; Wilson, R.; McCombie, W.R.

submitted to the EMBL Data Library, November 1998

A;Description: Sequence of A. thaliana BAC T10P11 from chromosome IV.

A;Reference number: Z14248

A;Accession: T01098

A;Status: translated from GB/EMBL/DDJB

A;Molecule type: DNA

A;Residues: I-1503 &lt;KAP&gt;

A;Cross-references: EMBL:AC002330; NID:g2262135; PID:g3892053

A;Experimental source: cultivar Columbia

R;Kalicki, J.; Elliott, G.; Cloud, J.

submitted to the EMBL Data Library, May 1998

A;Description: The sequence of A. thaliana T14P8.

A;Reference number: Z14290

A;Accession: T01299

A;Status: translated from GB/EMBL/DDJB

A;Molecule type: DNA

A;Residues: I-1503 &lt;KAP&gt;

A;Cross-references: EMBL:AF069298; NID:g3193282; PID:g3193301

A;Experimental source: cultivar Columbia

C;Genetics:

A;Map position: 4

A;Intron: 22/1

A;Note: T10P11.19- T14P8.24

C;Species: chloroplast; GTP binding; membrane protein; nucleotide binding; P-loop

F;862-B89/Region: nucleotide-binding motif A (P-loop)

Query Match 8.9%; Score 7; DB 2; Length 1503;

Best Local Similarity 100.0%; Pred. No. 53;

Matches 7; Conservative 0; Mismatches 0;

Indels 0; Gaps 0;

Qy 9 PAAVPVP 15

Db 1222 PAAVPVP 1228

RESULT 15

T37057 probable multi-domain beta keto-acyl synthase - *Streptomyces coelicolor*C;Species: *Streptomyces coelicolor*

C;Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 28-Jul-2000

C;Accession: T37057

R;Seeger, K.J.; Harris, D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.

submitted to the EMBL Data Library, August 1999

A;Reference number: Z21620

A;Accession: T37057

A;Status: preliminary; translated from GB/EMBL/DDJB

A;Molecule type: DNA

A;Residues: I-2240 &lt;SEE&gt;

A;Experimental source: strain A3(2)

A;Gene: SCOEDB:SCU21.08

C;Superfamily: 3-oxoacyl-[acyl-carrier-protein] synthase I homology

F;24-435/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology &lt;OAS&gt;

Query Match 8.9%; Score 7; DB 2; Length 2240;  
Best Local Similarity 100.0%; Pred. No. 76;

Matches 7; Conservative 0; Mismatches 0;

Indels 0; Gaps 0;

Qy 9 PAAVPVP 15

Db 1414 PAAVPVP 1420

Search completed: January 14, 2003, 17:14:59  
Job time : 7.6004 secs

Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 14, 2003, 17:08:59 ; Search time 10.1526 Seconds (without alignments)  
 1036.856 Million cell updates/sec

Title:	US-09-506-079E-1
Sequence:	1 GHSXXPRPAAVPVPKXQP. .... .VGRGXDPPDAHVAVLRSYEG 79
Scoring table:	OLIGO Gapext 60.0 , Gapext 60.0
Searched:	908470 seqs., 133250620 residues
Word size :	0
Total number of hits satisfying chosen parameters:	908470
Minimum DB seq length:	0
Maximum DB seq length:	200000000
Post-processing: Listing first 45 summaries	
Database :	A_Geneseq_101002:*
	1: /SIDS2/gcdata/geneseq/geneseq/geneseqp-emb1/AA1980.DAT:*
	2: /SIDS2/gcdata/geneseq/geneseqp-emb1/AA1981.DAT:*
	3: /SIDS2/gcdata/geneseq/geneseqp-emb1/AA1982.DAT:*
	4: /SIDS2/gcdata/geneseq/geneseqp-emb1/AA1983.DAT:*
	5: /SIDS2/gcdata/geneseq/geneseqp-emb1/AA1984.DAT:*
	6: /SIDS2/gcdata/geneseq/geneseqp-emb1/AA1985.DAT:*
	7: /SIDS2/gcdata/geneseq/geneseqp-emb1/AA1986.DAT:*
	8: /SIDS2/gcdata/geneseq/geneseqp-emb1/AA1987.DAT:*
	10: /SIDS2/gcdata/geneseq/geneseqp-emb1/AA1989.DAT:*
	11: /SIDS2/gcdata/geneseq/geneseqp-emb1/AA1990.DAT:*
	12: /SIDS2/gcdata/geneseq/geneseqp-emb1/AA1992.DAT:*
	13: /SIDS2/gcdata/geneseq/geneseqp-emb1/AA1993.DAT:*
	14: /SIDS2/gcdata/geneseq/geneseqp-emb1/AA1994.DAT:*
	15: /SIDS2/gcdata/geneseq/geneseqp-emb1/AA1995.DAT:*
	16: /SIDS2/gcdata/geneseq/geneseqp-emb1/AA1996.DAT:*
	17: /SIDS2/gcdata/geneseq/geneseqp-emb1/AA1997.DAT:*
	18: /SIDS2/gcdata/geneseq/geneseqp-emb1/AA1998.DAT:*
	19: /SIDS2/gcdata/geneseq/geneseqp-emb1/AA1999.DAT:*
	20: /SIDS2/gcdata/geneseq/geneseqp-emb1/AA2000.DAT:*
	21: /SIDS2/gcdata/geneseq/geneseqp-emb1/AA2001.DAT:*
	22: /SIDS2/gcdata/geneseq/geneseqp-emb1/AA2002.DAT:*
	23: /SIDS2/gcdata/geneseq/geneseqp-emb1/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No. ↴ Score ↴ Query length ↴ Match length DB ID ↴ Description

RESULT 1  
 ID AAY97239 standard; protein; 79 AA.  
 XX  
 AC AAY97239;  
 XX  
 DT 04-DEC-2000- (first entry)  
 DE HER-2 C-terminal polymorphic extracellular domain IIIa.  
 XX  
 KW HER-2; erbB-2; oncogene; receptor-like tyrosine kinase; insertion; extracellular domain IIIa; antagonist; intron 8; C-terminal extension; truncated HER-2; p68; dimerization inhibitor; cytosolic.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 2 /note= "Preferably Ser"  
 FT Misc-difference 5 /note= "Preferably Pro"  
 FT Misc-difference 6 /note= "Preferably Leu"  
 FT Misc-difference 16 /note= "Preferably Gln"  
 FT Misc-difference 18 /note= "Preferably Ile"  
 FT Misc-difference 21 /note= "Changes from glycine"  
 FT Misc-difference 36 /note= "Preferably Ile"  
 FT Misc-difference 54 /note= "Preferably Arg"

ALIGNMENTS

Result No.	Score	Query length	Match length	DB ID	Description
1	69	87.3	79	21 AAY97239	HER-2 C-terminal P
2	69	87.3	79	22 AAE09180	Human p68HER-2 ECD
3	69	87.3	79	23 AAE20347	Human HER2 intron
4	44	55.7	79	22 AAE09182	Human p68HER-2 ECD
5	41	51.9	419	22 AAE09181	Human p68HER-2 gen
6	41	51.9	419	22 AAE09210	Human p68HER-2 gen
7	41	51.9	419	22 AAE09211	Human p68HER-2 gen
8	41	51.9	419	22 AAE09212	Human p68HER-2 gen
9	41	51.9	419	22 AAE09213	Human p68HER-2 gen
10	41	51.9	419	23 AAE20348	Human truncated HE

FT Misc-difference 64 /note= "preferably leu"  
 FT Misc-difference 73 /note= "Preferably Asn"  
 XX  
 PN WO200044403-A1.  
 XX PD 03-AUG-2000.  
 XX PA 20-JAN-2000; 2000WO-US01484.  
 XX PR 20-JAN-1999; 99US-0234208.  
 XX PA (UYOR-) UNTV OREGON HEALTH SCI.  
 XX PI Doherty JK, Clinton GM, Adelman JP;  
 XX DR WPI; 2000-499287/44.

XX Using polypeptides and antibodies that bind to the extracellular domain of the receptor-like tyrosine kinase HER-2 to treat solid tumors of the breast, lung, ovaries and colon

PS Claim 1; Page 39; 46pp; English.

CC HER-2/neu (erbB-2) oncogene encodes a receptor-like tyrosine kinase. The extracellular domain of p185-HER-2 is proteolytically shed from breast carcinoma cells in culture and is found in serum of some cancer patients and may be a serum marker of metastatic breast cancer. An alternative CC mRNA of 4.8 kb with a 274 bp insert (intron 8) has been identified. The retained intron is in-frame and encodes a 79 amino acid extension designated ECDIIIA (the present sequence), which is inserted at residue 340 of p185-HER-2. The alternative mRNA predicts a truncated HER-2 protein (approximately 68 kDa) that lacks the transmembrane and intracellular domains (see AAY97240.). p68HER-2 specifically binds to p185-HER-2 without activating HER-2. It could therefore block dimerization of p185-HER-2. The p68HER-2 polypeptide binds to a site on the ECD of HER-2 that is different from the site of binding for Herceptin (RTM) (a marketed humanized monoclonal antibody that is used for the treatment of cancer and binds to the ECD of HER-2). The methods, compositions, polypeptides and antibodies are used to treat solid tumours such as breast cancer, small cell lung carcinoma, ovarian cancer and/or colon cancer, especially where over-expression of HER-2 is indicated.

XX Sequence 79 AA:

Query Match 87.3%; Score 69; DB 21; Length 79;  
 Best Local Similarity 100.0%; Pred. No. 3.2e-66; Indels 0; Gaps 0;  
 Matches 79; Conservative 0; Mismatches 0; Gaps 0;

QY 1 GKHGXKRPAVVPVKRQPKAPWPLSFLRSPWDYSAFISLPLAQLPLSPMSVX/SPVSV 60  
 Db 1 GKHGXKRPAVVPVKRQPKAPWPLSFLRSPWDYSAFISLPLAQLPLSPMSVX/SPVSV 60

QY 61 GRGXDPDAHVAVVLRLSRVEG 79  
 Db 61 GRGXDPDAHVAVVLRLSRVEG 79

RESULT 2

AEE09180 ID AAE09180 standard; peptide; 79 AA.  
 AC AAE09180;  
 XX  
 DT 15-NOV-2001 (first entry)  
 DE Human p68HER-2 ECDIIIA generic sequence #1.  
 KW HER-2; herstatin; antagonist; extracellular domain; ECD; Herceptin;  
 KW solid tumour; cancer; polymorphism; cytostatic; gene therapy;  
 KW p68HER-2; ECDIIIA.

XX OS Homo sapiens.  
 XX FH Key Location/qualifiers  
 FT Misc-difference 2 /label= Unknown  
 FT Misc-difference 5 /note= "Encoded by WCC"  
 FT Misc-difference 16 /label= Unknown  
 FT Misc-difference 17 /note= "Encoded by CYG"  
 FT Misc-difference 6 /label= Unknown  
 FT Misc-difference 18 /label= Unknown  
 FT Misc-difference 19 /note= "Encoded by CYC"  
 FT Misc-difference 20 /label= Unknown  
 FT Misc-difference 21 /label= Unknown  
 FT Misc-difference 22 /note= "Encoded by GNC"  
 FT Misc-difference 36 /label= Unknown  
 FT Misc-difference 37 /note= "Encoded by ATR"  
 FT Misc-difference 54 /label= Unknown  
 FT Misc-difference 55 /note= "Encoded by CST"  
 FT Misc-difference 64 /label= Unknown  
 FT Misc-difference 73 /label= Unknown  
 FT Misc-difference 74 /note= "Encoded by CYG"  
 FT Misc-difference 75 /label= Unknown  
 FT Misc-difference 76 /note= "Encoded by SAC"  
 FT WO20161356-A1.  
 XX PD 23-AUG-2001.  
 XX PF 16-FEB-2001; 2001WO-US05327.  
 XX PR 16-FEB-2000; 2000US-0506079.  
 XX PA (UYOR-) UNTV OREGON HEALTH SCI.  
 XX PI Clinton G, Henner WD, Evans A;  
 XX DR WPI; 2001-529934/58.  
 XX N-PSDB; AAD15844.  
 PT New polypeptide, which binds to the extracellular domain of HER-2 for the treatment of hard tumors

PS Claim 1; Page 52-53; 61pp; English.

CC The invention relates to novel HER-2 (herstatin-2) antagonist particularly a polypeptide that binds to the extracellular domain (ECD) of HER-2 at a site that is different from the binding site of humanised antibody, Herceptin, at an affinity of at least 10<sup>-8</sup>. The present invention is based upon the initial discovery of an alternative HER-2 mRNA transcript with 274 bp insert of intron 8. The translation product of the alternative transcript is a truncated HER-2 protein designated p68HER-2 which lacks the transmembrane and intracellular domains of p185HER-2 but contains ECD I, II of the p185HER-2 and the novel ECDIIIA. The ECDIIIA-containing polypeptides bind tightly to, and thus antagonise the HER-2 receptor. The peptides, which bind to an HER-2 ECD, and the nucleic acids encoding those are useful to treat, diagnose and identify solid tumours. The present sequence is human p68HER-2 ECDIIIA peptide generic sequence.

SO Sequence 79 AA;

Query Match 87.3%; Score 69; DB 22; Length 79;

			Best local Similarity 100.0%; Pred. No. 3.2e-66; Matches 79; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	DR N-PDB; AAD3238.
OY	1	GXHSXXPRPAAVPVPKXQPKPAHPVLSLRLPSPDXVSAYFSIPLAPISPLSPTSVXISPVSV 60	PT Treating solid tumor characterized by expression of endothelial growth factor receptor, involves administering recombinant herstatin that binds to extracellular domain of the endothelial growth factor receptor	
Db	1	GKHSXXPRPAAVPVPKXQPKPAHPVLSLRLPSPDXVSAYFSIPLAPISPLSPTSVXISPVSV 60	PT	
RESULT	3	61 GRGKDPAHVAVALSRYEG 79	XX	
AAR0347			PS Claim 1: Page 77; 82pp; English.	
ID	AAE20347	standard; Protein; 79 AA.	XX	
XX			CC The present invention relates to a method for treating a solid tumour characterised by endothelial growth factor receptor (EGFR) expression. The method involves administering an agent that binds to an extracellular domain (ECD) of EGFR. The invention also relates to a naturally occurring inhibitor of HER-2 receptor tyrosine kinase called herstatin. The co-expression of herstatin with p185HER2 causes a striking reduction in cell growth that corresponds with suppression of p185 autophosphorylation. The method or a pharmaceutical composition is useful for treating a solid tumour (selected from squamous cell carcinoma, lung carcinoma, colon carcinoma and glial cell tumour) characterised by EGFR expression. The present sequence is a protein encoded by human HER2 intron 8 polymorphic sequence.	
AC	AAE20347;		CC	
XX			CC	
DT	18-JUN-2002	(first entry)	CC	
DE	Human HER2 intron 8 polymorphic sequence encoded protein.		CC	
XX			CC	
KW	Human; tumour; endothelial growth factor receptor; EGFR; cytostatic; hestatin; HER-2 receptor tyrosine kinase; squamous cell carcinoma; lung; colon; glial cell tumour; cell growth; polymorphism.		CC	
KW	herstatin; HER-2 receptor tyrosine kinase; squamous cell carcinoma; lung; colon; glial cell tumour; cell growth; polymorphism.		CC	
OS	Homo sapiens.		CC	
XX			CC	
Key		Location/Qualifiers	XX	
FH	Misc-difference 2	/label= "Encoded by CYG"	Sequence 79 AA;	
FT		/note= "Encoded by WCC"		
FT	Misc-difference 5	/label= "Leu, Pro		
FT		/note= "Encoded by CYG"		
FT	Misc-difference 6	/label= "Pro, Leu		
FT		/note= "Encoded by CYG"		
FT	Misc-difference 16	/label= "Leu, Gln		
FT		/note= "Encoded by CWG"		
FT	Misc-difference 18	/label= "Met, Leu		
FT		/note= "Encoded by ATR"		
FT	Misc-difference 21	/label= "GLY, Ala, Val		
FT		/note= "Encoded by GNC"		
FT	Misc-difference 35	/label= "Leu, Ile		
FT		/note= "Encoded by MTA"		
FT	Misc-difference 54	/label= "Pro, Arg		
FT		/note= "Encoded by CSR"		
FT	Misc-difference 64	/label= "Pro, Leu		
FT		/note= "Encoded by CYG"		
FT	Misc-difference 73	/label= "Asp, Asn		
FT		/note= "Encoded by SAC"		
XX	WO200214470-A2.			
XX	21-FEB-2002.			
XX	14-AUG-2001; 2001WO-US25502.			
XX	14-AUG-2000; 2000US-0638834.			
PA	(UWOR-) UNIV OREGON HEALTH SCI.			
XX	Clinton GM;			
PI	Clinton GM;			
XX	WPI; 2002-269185/31.			

FT /label= Unknown  
 FT /note= "Encoded by GNC"  
 FT Misc-difference 31 /label= Unknown  
 FT /note= "Encoded by AKA"  
 FT Misc-difference 36 /label= Unknown  
 FT /note= "Encoded by MTA"  
 FT Misc-difference 54 /label= Unknown  
 FT /note= "Encoded by CST"  
 FT Misc-difference 64 /label= Unknown  
 FT /note= "Encoded by CYG"  
 FT Misc-difference 73 /label= Unknown  
 FT /note= "Encoded by VAC"  
 PN WO200161356-A1.  
 PD 23-AUG-2001.  
 XX PR 16-FEB-2001; 2001WO-US05327.  
 XX PR 16-FEB-2000; 2000US-0506079.  
 PA (UYOR-) UNIV OREGON HEALTH SCI.  
 PI Clinton G., Henner WD, Evans A;  
 XX DR WPI; 2001-529934/58.  
 DR N-PSDB; AAD15852.  
 XX PT New polypeptide, which binds to the extracellular domain of HER-2 for  
 PT the treatment of hard tumors -  
 XX PS Claim 1; Page 57; 61pp; English.  
 CC The invention relates to novel HER-2 (herstatin-2) antagonist  
 CC particularly a polypeptide that binds to the extracellular domain (ECD)  
 CC of HER-2 at a site that is different from the binding site of humanised  
 antibody, Herceptin, at an affinity of at least 10<sup>-8</sup>. The present  
 CC invention is based upon the initial discovery of an alternative HER-2  
 CC mRNA transcript with 274 bp insert of intron 8. The translational product  
 CC of the alternative transcript is a truncated HER-2 protein designated  
 CC p68HER-2 which lacks the transmembrane and intracellular domains of  
 CC the p185HER-2 but contains ECD I, II of the p185HER-2 and the novel ECDIIIA.  
 CC The ECDIIIA-containing polypeptides bind tightly to, and thus antagonise  
 CC the HER-2 receptor. The peptides, which bind to an HER-2 ECD, and the  
 CC nucleic acids encoding these are useful to treat, diagnose and identify  
 CC solid tumours. The present sequence is human p68HER-2 ECDIIA  
 CC peptide generic sequence.  
 XX SQ Sequence 79 AA;  
 Query Match 55.7%; Score 44;  
 Best local Similarity 100.0%; Pred. No. 1.9e-39;  
 Matches 48; Conservative 0; Mishatches 0; Indels 0; Gaps 0;  
 Oy 32 PSWDKXSAFYSPLAPLSPSTSXVISPVSVGKDPDAHVAYKLSRYEG 79  
 Db 32 PSWDKXSAFYSPLAPLSPSTSXVISPVSVGKDPDAHVAYKLSRYEG 79  
 RESULT 5  
 AAE09181 ID AAE09181 standard; Protein; 419 AA.  
 AC XX AAE09181;  
 XX DT 15-NOV-2001 (first entry)  
 XX DE Human p68HER-2 generic sequence #1.  
 XX XX KW HER-2; herstatin; antagonist; extracellular domain; ECD; Herceptin;  
 KW solid tumour; cancer; polymorphism; cytostatic; gene therapy;  
 KW p68HER-2; ECDIIIA.  
 OS Homo sapiens.  
 XX FH Key Location/Qualifiers  
 FT Region 1..340  
 FT /note= "Identical to N-terminal region of p185HER-2"  
 FT Domain 341..419  
 FT /label= ECDIIIA  
 FT Misc-difference 124 /note= "Represented as Agn in the sequence shown in  
 FT the specification"  
 FT Misc-difference 125 /note= "Represented as Agn in the sequence shown in  
 FT the specification"  
 FT Misc-difference 342 /label= Unknown  
 FT /note= "Encoded by WCC"  
 FT Misc-difference 346 /label= Unknown  
 FT /note= "Encoded by CYG"  
 FT Misc-difference 345 /label= Unknown  
 FT /note= "Encoded by CYC"  
 FT Misc-difference 356 /label= Unknown  
 FT /note= "Encoded by CWG"  
 FT Misc-difference 358 /label= Unknown  
 FT /note= "Encoded by ATR"  
 FT Misc-difference 361 /label= Unknown  
 FT /note= "Encoded by GNC"  
 FT Misc-difference 376 /label= Unknown  
 FT /note= "Encoded by MTA"  
 FT Misc-difference 389 /note= "Encoded by AGC"  
 FT Misc-difference 394 /label= Unknown  
 FT /note= "Encoded by CST"  
 FT Misc-difference 404 /label= Unknown  
 FT /note= "Encoded by CYG"  
 FT Misc-difference 413 /label= Unknown  
 FT /note= "Encoded by SAC"  
 PN WO200161356-A1.  
 XX PD 23-AUG-2001.  
 XX PR 16-FEB-2001; 2001WO-US05327.  
 XX PR 16-FEB-2000; 2000US-0506079.  
 PA (UYOR-) UNIV OREGON HEALTH SCI.  
 PI Clinton G., Henner WD, Evans A;  
 XX DR WPI; 2001-529934/58.  
 DR N-PSDB; AAD15844.  
 XX PT New polypeptide, which binds to the extracellular domain of HER-2 for  
 PT the treatment of hard tumors -  
 XX PS Claim 8; Page 53-54; 61pp; English.  
 XX CC The invention relates to novel HER-2 (herstatin-2) antagonist





CC antibody, Herceptin, at an affinity of at least 10<sup>8</sup>. The present invention is based upon the initial discovery of an alternative HER-2 mRNA transcript with 274 bp insert of intron 8. The translation product of the alternative transcript is a truncated HER-2 protein designated p68HER-2 which lacks the transmembrane and intracellular domains of p185HER-2 but contains ECD I, II of the p185HER-2 and the novel ECDIIA. The ECDIIA-containing polypeptides bind tightly to, and thus antagonise the HER-2 receptor. The peptides, which bind to an HER-2 ECD, and the nucleic acids encoding these are useful to treat, diagnose and identify solid tumours. The present sequence is human p68HER-2 generic protein containing ECDIIA variant sequence.

Note: The present sequence is not shown in the specification but is derived from p68HER-2 generic sequence (SEQ ID NO:2) shown in the sequence listing (AAE09181).

CC sequence 419 AA;

Query Match 51.9%; Score 41; DB 22; Length 419;  
Best Local Similarity 100.0%; Pred. No. 1.5e-35; Indels 0; Gaps 0;  
Matches 48; Conservative 0; Mismatches 0; XX

OY 1 GXHSXKPRPAVPPVXRXQXPXPAHVLSTLRPSMDXVSAYFSLPLAPL 48  
Db 341 GXHSXKPRPAVPPVXRXQXPXPAHVLSTLRPSMDXVSAYFSLPLAPL 388

RESULT 9  
AAE09212  
ID AAE09212 standard; Protein; 419 AA.  
XX  
AC AAE09212;  
XX  
DT 15 NOV-2001 (first entry)  
XX  
DE Human p68HER-2 generic protein variant 10.  
XX  
HER-2; herstatin; antagonist; extracellular domain; ECD; Herceptin;  
KW solid tumour; cancer; polymorphism; cytostatic; gene therapy;  
KW p68HER-2; ECDIIA; variant.  
XX  
OS Homo sapiens.  
XX

FH Key Location/Qualifiers  
FT Region  
FT Domain  
FT Label= ECDIIA\_variant  
FT /note= "Extracellular domain IIIA variant"  
FT Misc-difference 124  
FT /note= "Represented as Agn in the parent sequence shown  
in the specification"  
FT Misc-difference 125  
FT /note= "Represented as Agn in the parent sequence shown  
in the specification"  
FT Misc-difference 342  
FT /label= Unknown  
FT Misc-difference 345  
FT /label= Unknown  
FT Misc-difference 346  
FT /label= Unknown  
FT Misc-difference 356  
FT /label= Unknown  
FT Misc-difference 358  
FT /label= Unknown  
FT Misc-difference 361  
FT /label= Unknown  
FT Misc-difference 376  
FT /label= Unknown  
FT Misc-difference 394  
FT /label= Unknown  
FT Misc-difference 404  
FT /label= Unknown

The invention relates to novel HER-2 (herstatin-2) antagonist particularly a polypeptide that binds to the extracellular domain (ECD) of HER-2 at a site that is different from the binding site of humanised antibody, Herceptin, at an affinity of at least 10<sup>8</sup>. The present invention is based upon the initial discovery of an alternative HER-2 mRNA transcript with 274 bp insert of intron 8. The translation product of the alternative transcript is a truncated HER-2 protein designated p68HER-2 which lacks the transmembrane and intracellular domains of p185HER-2 but contains ECD I, II of the p185HER-2 and the novel ECDIIA. The ECDIIA-containing polypeptides bind tightly to, and thus antagonise the HER-2 receptor. The peptides, which bind to an HER-2 ECD, and the nucleic acids encoding these are useful to treat, diagnose and identify solid tumours. The present sequence is human p68HER-2 generic protein containing ECDIIA variant sequence.

Note: The present sequence is not shown in the specification but is derived from p68HER-2 generic sequence (SEQ ID NO:2) shown in the sequence listing (AAE09181).

SQ Sequence 419 AA;  
XX  
PT New polypeptide, which binds to the extracellular domain of HER-2 for the treatment of hard tumors -  
XX  
PS Example 11; Page 419; English.

Query Match 51.9%; Score 41; DB 22; Length 419;  
Best Local Similarity 100.0%; Pred. No. 1.5e-35; Indels 0; Gaps 0;  
Matches 48; Conservative 0; Mismatches 0; XX

OY 1 GXHSXKPRPAVPPVXRXQXPXPAHVLSTLRPSMDXVSAYFSLPLAPL 48  
Db 341 GXHSXKPRPAVPPVXRXQXPXPAHVLSTLRPSMDXVSAYFSLPLAPL 388

RESULT 10  
AAE0348  
ID AAE0348 standard; Protein; 419 AA.  
XX  
AC AAE0348;  
XX  
DT 18-JUN-2002 (first entry)  
XX  
DE Human truncated HER2 protein.  
XX  
Human; tumour; endothelial growth factor receptor; EGFR; cytostatic; herstatin; HER-2 receptor tyrosine kinase; squamous cell carcinoma; lung; colon; gial cell tumour; cell growth.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT Misc-difference 342  
FT /label= Thr, Ser  
FT Misc-difference 345  
FT /label= Leu, Pro  
FT Misc-difference 346

FT /note= "p68HER-2 generic sequence (AAE09181) xaa substituted with Asn"  
FT  
XX  
PN WO200161356-A1.  
XX  
PD 23-AUG-2001.  
XX  
PR 16-FEB-2001; 2001WO-US03327.  
XX  
PR 16-FEB-2000; 2000US-050079.  
XX  
PA (UYOR-) UNIV OREGON HEALTH SCI.  
XX  
DR Clinton G, Henner WD, Evans A;  
XX  
WPI: 2001-529334/5B.

KW truncated HER-2; p68; dimerization inhibitor; cytostatic.  
 XX  
 FT  
 FT /label= pro, Leu  
 FT /label= Leu, Gln  
 XX Homo sapiens.  
 XX  
 Key Location/Qualifiers  
 FH  
 FT  
 FT /note= "represented as Arg"  
 FT  
 FT /note= "represented as Arg"  
 FT  
 FT /note= "Preferably Ser"  
 FT  
 FT /note= "Preferably Pro"  
 FT  
 FT /note= "Preferably Ile"  
 FT  
 FT /note= "Preferably Gln"  
 FT  
 FT /note= "Preferably Ieu"  
 FT  
 FT /note= "changes from glycine"  
 FT  
 FT /note= "Preferably Ile"  
 FT  
 FT /note= "Preferably Arg"  
 FT  
 FT /note= "Preferably Ieu"  
 FT  
 FT /note= "Preferably Asn"  
 XX  
 PD 21-FEB-2002.  
 XX  
 PF 14-AUG-2001; 2001WO-US25502.  
 XX  
 PR 14-AUG-2000; 2000US-0638834.  
 XX  
 PA (UYCR-) UNIV OREGON HEALTH SCI.  
 PI Clinton GM;  
 XX  
 DR WPI; 2002-269185/31.  
 XX  
 PT Treating solid tumor characterized by expression of endothelial growth factor receptor, involves administering recombinant herstatin that binds to extracellular domain of the endothelial growth factor receptor  
 PT  
 PS Claim 1; Page 78-80; 82pp; English.  
 XX  
 CC The present invention relates to a method for treating a solid tumour characterised by endothelial growth factor receptor (EGFR) expression. The method involves administering an agent that binds to an extracellular domain (ECD) of EGFR. The invention also relates to a naturally occurring inhibitor of HER-2 receptor tyrosine kinase called herstatin. The co-expression of herstatin with p185HER2 causes a striking reduction in cell growth that corresponds with suppression of p185 autophosphorylation. The method or a pharmaceutical composition is useful for treating a solid tumour (selected from squamous cell carcinoma, lung carcinoma, colon carcinoma and gliial cell tumour) characterised by EGFR expression. The present sequence is human truncated HER2 protein that lacks transmembrane and intracellular domains.  
 XX Sequence 419 AA:  
 SQ Query Match 51.9%; Score 41; DB 23; Length 419;  
 Best Local Similarity 100.0%; Pred. No. 1.5e-35;  
 Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 1 GXSXXPRPAAVFVPPXKXQXPXPAHPVLSFLPSPWDXYSAFYSLPLAPL 48  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 341 GXHSXXPRPAAVFPXKXQXPXPAHPVLSFLPSPWDXYSAFYSLPLAPL 388  
 RESULT 11  
 AAY9740  
 ID AAY9740 standard; protein: 420 AA.  
 XX  
 AC AAY9740;  
 XX 04-DEC-2000 (first entry)  
 DE Truncated HER-2, p68 HER-2.  
 KW HER-2; erbB-2; oncogene; receptor-like tyrosine kinase; insertion; extracellular domain IIIa; antagonist; intron 8; C-terminal extension;

KW truncated HER-2; p68; dimerization inhibitor; cytostatic.  
 XX  
 FT  
 FT /label= pro, Leu  
 FT /label= Leu, Gln  
 XX Homo sapiens.  
 XX  
 Key Location/Qualifiers  
 FH  
 FT  
 FT /note= "represented as Arg"  
 FT  
 FT /note= "represented as Arg"  
 FT  
 FT /note= "Preferably Ser"  
 FT  
 FT /note= "Preferably Pro"  
 FT  
 FT /note= "Preferably Ile"  
 FT  
 FT /note= "Preferably Gln"  
 FT  
 FT /note= "Preferably Ieu"  
 FT  
 FT /note= "changes from glycine"  
 FT  
 FT /note= "Preferably Ile"  
 FT  
 FT /note= "Preferably Arg"  
 FT  
 FT /note= "Preferably Ieu"  
 FT  
 FT /note= "Preferably Asn"  
 XX  
 PD WO20044403-A1.  
 XX  
 PF 20-JAN-1999; 99US-0234208.  
 XX  
 PA (UYCR-) UNIV OREGON HEALTH SCI.  
 XX  
 PI Doherty JK, Clinton GM, Adelman JP;  
 XX  
 DR WPI; 2000-49928744.  
 XX  
 PT Using polypeptides and antibodies that bind to the extracellular domain of the receptor-like tyrosine kinase HER-2 to treat solid tumors of the breast, lung, ovaries and colon  
 XX  
 PS WO20004403-A1.  
 XX  
 PR 03-AUG-2000.  
 XX  
 PF 20-JAN-2000; 2000WO-US01484.  
 XX  
 PR 20-JAN-1999; 99US-0234208.  
 XX  
 PA (UYCR-) UNIV OREGON HEALTH SCI.  
 XX  
 PI Doherty JK, Clinton GM, Adelman JP;  
 XX  
 DR WPI; 2000-49928744.  
 XX  
 PT Using polypeptides and antibodies that bind to the extracellular domain of the receptor-like tyrosine kinase HER-2 to treat solid tumors of the breast, lung, ovaries and colon  
 XX  
 PS WO20004403-A1.  
 XX  
 CC HER-2/neu (erbB-2) oncogene encodes a receptor-like tyrosine kinase. The extracellular domain of p185-HER-2 is proteolytically shed from breast carcinoma cells in culture and is found in serum of some cancer patients and may be a serum marker of metastatic breast cancer. An alternative HER-2 mRNA of 4.8 kb with a 274 bp in-frame insert (intron 8) has been identified. The retained intron 8 encodes a 79 amino acid extension designated ECDIIA (the present sequence), which is inserted at residue 340 of p185-HER-2. The alternative mRNA predicts a truncated HER-2 protein (approximately 68 kDa) that lacks the transmembrane and intracellular domains (see AAY9740). p68HER-2 specifically binds to p185-HER-2 without activating HER-2. It could therefore block dimerization of p185-HER-2. The p68-HER-2 polypeptide binds to a site on the ECD of HER-2 that is different from the site of binding for Herceptin (RMM) (a marketed humanized monoclonal antibody that is used for the treatment of cancer and binds to the ECD of HER-2). The methods, compositions, polypeptides and antibodies are used to treat solid tumours such as breast cancer, small cell lung carcinoma, ovarian cancer and/or colon cancer, especially where over-expression of HER-2 is indicated.  
 XX  
 Sequence 420 AA:  
 SQ Query Match 51.9%; Score 41; DB 21; Length 420;

	PT	the treatment of hard tumors -
	XX	
OY	PS	Example 11; Page -; 61pp; English.
1	XX	
GXHSXXPRAVPVPPXRXQXPXAHPLSLRPSWDXSAFVSLPLAPL 48	CC	The invention relates to novel HER-2 (herstatin-2) antagonist particularly a polypeptide that binds to the extracellular domain (ECD) of HER-2 at a site that is different from the binding site of humanised antibody, Herceptin, at an affinity of at least 10 <sup>8</sup> . The present invention is based upon the initial discovery of an alternative HER-2 mRNA transcript with 274 bp insert of intron 8. The translation product of the alternative transcript is a truncated HER-2 protein designated p68HER-2 which lacks the transmembrane and intracellular domains of p185HER-2 but contains ECD I, II of the p185HER-2 and the novel ECDIIa. The ECDIIa-containing polypeptides bind tightly to, and thus antagonise the HER-2 receptor. The peptides, which bind to an HER-2 ECD, and the nucleic acids encoding these are useful to treat, diagnose and identify solid tumours. The present sequence is human p68HER-2 generic protein containing ECDIIa variant sequence.
342	CC	Note: The present sequence is not shown in the specification but is derived from p68HER-2 generic sequence (SEQ ID NO:2) shown in the sequence listing (AAE09181).
GXHSXXPRAVPVPPXRXQXPXAHPLSLRPSWDXSAFVSLPLAPL 389	CC	CC
Db	CC	CC
AAE09203	CC	CC
ID	CC	CC
AAE09203 standard; Protein; 419 AA.	CC	CC
XX	CC	CC
AC	CC	CC
AAE09203;	CC	CC
XX	CC	CC
15 -NOV-2001 (first entry)	CC	CC
DE	CC	CC
Human p68HER-2 generic protein variant 1.	CC	CC
XX	CC	CC
KW	CC	CC
HER-2; herstatin; antagonist; extracellular domain; ECD; Herceptin;	CC	CC
KW	CC	CC
solid tumour; cancer; polymorphism; cytostatic; gene therapy;	CC	CC
KW	CC	CC
p68HER-2; ECDIIa; variant.	CC	CC
XX	CC	CC
OS	CC	CC
Homo sapiens.	CC	CC
XX	CC	CC
FH	Sequence	419 AA;
Key	Query Match	50 %;
Region	Best Local Similarity	100.0%;
/note= "Identical to N-terminal region of p185HER-2"	Score	40;
341..419	DB	22;
Domain	Matches	Length 419;
/label= ECDIIa_variant	ID	No. 1.8e-34;
FT	AAE09204	Indels
/note= "Extracellular domain IIIa variant"	AC	0;
FT	AAE09204	Gaps
Misc-difference	XX	0;
124	DT	
/note= "Represented as Agn in the parent sequence shown	XX	
FT	15 -NOV-2001 (first entry)	
Misc-difference	XX	
125	DE	
/note= "Represented as Agn in the parent sequence shown	XX	
FT	Human p68HER-2 generic protein variant 2.	
Misc-difference	XX	
342	XX	
/note= "p68HER-2 generic sequence (AAE09181) Xaa	XX	
FT	XX	
Misc-difference	XX	
345	XX	
/label= Unknown	XX	
FT	XX	
Misc-difference	XX	
346	XX	
/label= Unknown	XX	
FT	XX	
Misc-difference	XX	
356	XX	
/label= Unknown	XX	
FT	XX	
Misc-difference	XX	
358	XX	
/label= Unknown	XX	
FT	XX	
Misc-difference	XX	
361	XX	
/label= Unknown	XX	
FT	XX	
Misc-difference	XX	
376	XX	
/label= Unknown	XX	
FT	XX	
Misc-difference	XX	
394	XX	
/label= Unknown	XX	
FT	XX	
Misc-difference	XX	
404	XX	
/label= Unknown	XX	
FT	XX	
Misc-difference	XX	
413	XX	
/label= Unknown	XX	
FT	XX	
XX	XX	
WO200161356-A1.	XX	
XX	XX	
PD	XX	
23 -AUG-2001.	XX	
XX	XX	
PF	XX	
16 -FEB-2001; 2001WO-US05327.	XX	
XX	XX	
PR	XX	
16 -FEB-2000; 2000US-0506079.	XX	
XX	XX	
(UYOR-) UNIV OREGON HEALTH SCI.	PA	
Clinton G, Henner WD, Evans A;	PI	
XX	XX	
DR; 2001-529934/58.	DR	
XX	XX	
• New polypeptide, which binds to the extracellular domain of HER-2 for	PT	

FT Misc-difference /label= Unknown  
 FT 376  
 FT /label= Unknown  
 FT 394  
 FT Misc-difference /label= Unknown  
 FT 404  
 FT Misc-difference /label= Unknown  
 FT 413  
 FT Misc-difference /label= Unknown  
 FT 413  
 PN WO200161356-A1.  
 XX PD 23-AUG-2001.  
 XX PF 16-FEB-2001; 2001WO-US05327.  
 PR 16-FEB-2000; 2000US-0506079.  
 XX PA (UYOR-) UNIV OREGON HEALTH SCI.  
 PI Clinton G, Henner WD, Evans A;  
 DR WPI; 2001-529934/58.  
 XX PT New polypeptide, which binds to the extracellular domain of HER-2 for  
 PT the treatment of hard tumors.  
 XX PS Example 11; Page - ; 61pp; English.  
 XX CC The invention relates to novel HER-2 (herstatin-2) antagonist  
 CC particularly a polypeptide that binds to the extracellular domain (ECD)  
 CC of HER-2 at a site that is different from the binding site of humanised  
 CC antibody, Herceptin, at an affinity of at least 10<sup>-8</sup>. The present  
 CC invention is based upon the initial discovery of an alternative HER-2  
 CC mRNA transcript with 274 bp insert of intron 8. The transcript product  
 CC of the alternative transcript is a truncated HER-2 protein designated  
 CC p68HER-2 which lacks the transmembrane and intracellular domains of  
 CC p185HER-2 but contains ECD I, II of the p185HER-2 and the novel ECDIIIA.  
 CC The ECDIIIA-containing polypeptides bind tightly to, and thus antagonise  
 CC the HER-2 receptor. The peptides, which bind to an HER-2 ECD, and the  
 CC nucleic acids encoding these are useful to treat, diagnose and identify  
 CC solid tumours. The present sequence is human p68HER-2 generic protein  
 CC containing ECDIIIA variant sequence.  
 CC Note: The present sequence is not shown in the specification but is  
 CC derived from p68HER-2 generic sequence (SEQ ID NO:2) shown in the  
 CC sequence listing (AAE09181).  
 XX SQ sequence 419 AA;

Query Match 48.1%; Score 38; DB 22; Length 419;  
 Best Local Similarity 100.0%; Pred. No. 2.4e-32; Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 7 PRPAAVPVPKXQPKXAHPVLSFLRPSWDXYSAFISIPLAPL 48  
 Db 347 PRPAAVPVPKXQPKXAHPVLSFLRPSWDXYSAFISIPLAPL 388

RESULT 14  
 AAE0205  
 ID AAE09205 standard; Protein: 419 AA.  
 XX AC AAE09205;  
 XX DT 15-NOV-2001 (first entry)  
 DE Human p68HER-2 generic protein variant 3.  
 KW HER-2; herstatin; antagonist; extracellular domain; ECD; Herceptin;  
 KW solid tumour; cancer; polymorphism; cytostatic; gene therapy;  
 KW p68HER-2; ECDIIIA; variant.  
 XX OS Homo sapiens.

XX Key location/qualifiers  
 FH Region 1..340  
 FT /note= "Identical to N-terminal region of p185HER-2"  
 FT Domain 341..419  
 FT /note= "Extracellular domain IIIA variant"  
 FT Misc-difference 124  
 FT /note= "Represented as Agn in the parent sequence shown  
 FT in the specification"  
 FT Misc-difference 125  
 FT /note= "Represented as Agn in the parent sequence shown  
 FT in the specification"  
 FT Misc-difference 342  
 FT /note= "Substituted with Leu"  
 FT Misc-difference 345  
 FT /note= "Label= Unknown  
 FT Misc-difference 346  
 FT /note= "Label= Agn in the specification"  
 FT Misc-difference 356  
 FT /note= "Label= Unknown  
 FT Misc-difference 358  
 FT /note= "Label= Unknown  
 FT Misc-difference 361  
 FT /note= "Label= Unknown  
 FT Misc-difference 376  
 FT /note= "Label= Unknown  
 FT Misc-difference 394  
 FT /note= "Label= Unknown  
 FT Misc-difference 404  
 FT /note= "Label= Unknown  
 FT Misc-difference 413  
 FT /note= "Label= Unknown  
 PN WO200161356-A1.  
 XX PD 23-AUG-2001.  
 XX PF 16-FEB-2001; 2001WO-US05327.  
 PR 16-FEB-2000; 2000US-0506079.  
 XX PA (UYOR-) UNIV OREGON HEALTH SCI.  
 PI Clinton G, Henner WD, Evans A;  
 DR WPI; 2001-529934/58.  
 XX PS Example 11; Page - ; 61pp; English.  
 XX CC The invention relates to novel HER-2 (herstatin-2) antagonist  
 CC particularly a polypeptide that binds to the extracellular domain (ECD)  
 CC of HER-2 at a site that is different from the binding site of humanised  
 CC antibody, Herceptin, at an affinity of at least 10<sup>-8</sup>. The present  
 CC invention is based upon the initial discovery of an alternative HER-2  
 CC mRNA transcript with 274 bp insert of intron 8. The transcript product  
 CC of the alternative transcript is a truncated HER-2 protein designated  
 CC p68HER-2 which lacks the transmembrane and intracellular domains of  
 CC p185HER-2 but contains ECD I, II of the p185HER-2 and the novel ECDIIIA.  
 CC The ECDIIIA-containing polypeptides bind tightly to, and thus antagonise  
 CC the HER-2 receptor. The peptides, which bind to an HER-2 ECD, and the  
 CC nucleic acids encoding these are useful to treat, diagnose and identify  
 CC solid tumours. The present sequence is human p68HER-2 generic protein  
 CC containing ECDIIIA variant sequence.  
 CC Note: The present sequence is not shown in the specification but is  
 CC derived from p68HER-2 generic sequence (SEQ ID NO:2) shown in the  
 CC sequence listing (AAE09181).  
 XX SQ Sequence 419 AA;

Query Match 48.1%; Score 38; DB 22; Length 419;  
 Best Local Similarity 100.0%; Pred. No. 2.4e-32; Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0; OS Homo sapiens.

DR WPI; 2001-529934/58.  
 XX PT New polypeptide, which binds to the extracellular domain of HER-2 for  
 PT the treatment of hard tumors.

Example 11: Page -; 61pp; English.

Query 7 PRPAAVPVPRXQXPAPAHVPLSFLRPSWDVSAFVSLPLAPL 48  
 DB 347 PRPAAVPVPRXQXPAPAHVPLSFLRPSWDVSAFVSLPLAPL 388

RES15 AAE09213  
 ID AAE09213 standard; Protein; 419 AA.  
 XX AC AAE09213;  
 XX DT 15-NOV-2001 (first entry)  
 XX DE Human p68HER-2 generic protein variant 11.  
 XX KW HER-2; heparstatin; antagonist; extracellular domain; ECD; Herceptin;  
 KW solid tumour; cancer; polymorphism; cytostatic; gene therapy;  
 KW p68HER-2; ECDiria; variant.  
 XX OS Homo sapiens.

Key Location/Qualifiers  
 FH Region 1..340  
 FT /note= "Identical to N-terminal region of p185HER-2"  
 FT Domain 341..419  
 FT /note= "Extracellular domain IIIa variant"  
 FT Misc-difference 124  
 FT /note= "Represented as Agn in the parent sequence shown  
 in the specification"  
 FT Misc-difference 125  
 FT /note= "Represented as Agn in the parent sequence shown  
 in the specification"  
 FT Misc-difference 342  
 FT /label= Unknown  
 FT Misc-difference 345  
 FT /label= Unknown  
 FT Misc-difference 346  
 FT /note= "p68HER-2 generic sequence (AAE09181) xaa  
 substituted with Ieu"  
 FT Misc-difference 356  
 FT /label= Unknown  
 FT Misc-difference 358  
 FT /label= Unknown  
 FT Misc-difference 361  
 FT /label= Unknown  
 FT Misc-difference 376  
 FT /label= Unknown  
 FT Misc-difference 394  
 FT /label= Unknown  
 FT Misc-difference 404  
 FT /label= Unknown  
 FT Misc-difference 413  
 FT /note= "p68HER-2 generic sequence (AAE09181) xaa  
 substituted with Asn"  
 XX WO200161356-A1.  
 XX PR 23-AUG-2001.  
 XX PD 16-FEB-2001; 2001WO-US05327.  
 XX PR 16-FEB-2000; 2000US-0506079.  
 XX PA (UYOR-) UNIV OREGON HEALTH SCI.  
 XX PI Clinton G, Henner WD, Evans A;

Query Match 48.1%; Score 38; DB 22; Length 419;  
 Best Local Similarity 100.0%; Pred. No. 2.4e-32; Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0; OS Homo sapiens.

DR WPI; 2001-529934/58.  
 XX PT New polypeptide, which binds to the extracellular domain of HER-2 for  
 PT the treatment of hard tumors.

The invention relates to novel HER-2 (herstatin-2) antagonist particularly a polypeptide that binds to the extracellular domain (ECD) of HER-2 at a site that is different from the binding site of humanised antibody, Herceptin, at an affinity of at least 10<sup>8</sup>. The present invention is based upon the initial discovery of an alternative HER-2 mRNA transcript with 274 bp insert of intron 8. The translation product of the alternative transcript is a truncated HER-2 protein designated p68HER-2 which lacks the transmembrane and intracellular domains of p185HER-2 but contains ECD I, II of the p185HER-2 and the novel ECDiria. The ECDiria-containing polypeptides bind tightly to, and thus antagonise the HER-2 receptor. The peptides, which bind to an HER-2 ECD, and the nucleic acids encoding these are useful to treat, diagnose and identify solid tumours. The present sequence is human p68HER-2 generic protein containing ECDiria variant sequence.

Note: The present sequence is not shown in the specification but is derived from p68HER-2 generic sequence (SEQ ID NO:2) shown in the sequence listing (AAE09181).

Sequence 419 AA;

Query Match 48.1%; Score 38; DB 22; Length 419;  
 Best Local Similarity 100.0%; Pred. No. 2.4e-32; Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0; OS Homo sapiens.

Query 7 PRPAAVPVPRXQXPAPAHVPLSFLRPSWDVSAFVSLPLAPL 48  
 DB 347 PRPAAVPVPRXQXPAPAHVPLSFLRPSWDVSAFVSLPLAPL 388

Search completed: January 14, 2003, 17:12:54  
 Job time : 11.1526 secs

**THIS PAGE BLANK (USPTO)**

Copyright (c) 1993 - 2003 Compugen Ltd.									
OM protein - protein search, using sw model									
Run on: January 14, 2003, 17:10:35 ; search time 14:3032 Seconds									
(without alignments)									
1215.012 Million cell updates/sec									
Title: US-09-506-079-E-2	Sequence: 1 MELAALCNGLLIALIPGA.....VGRGKXDPDAHVAVALRSYEG 419	Scoring table: Oligo Gapp 60.0 , Gapext 60.0	Word size : 0	Searched: 112892 seqs, 41476328 residues	Total number of hits satisfying chosen parameters: 112892	Minimum DB seq length: 0	Maximum DB seq length: 200000000	Post-processing: Listing first 45 summaries	Database: SwissProt_40:*
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.									
SUMMARIES									
Result No.	Scored	Query	Length	DB ID	Description	RT	RT	RT	ALIGNMENTS
1	83	19 8	1255	1 EBB2_HUMAN	P04626 homo sapien	ERBB2_HUMAN	STANDARD;	PRN; 1255 AA.	
2	51	2 2	1252	1 ERBB2_MSAU	Q66533 mesocricetus auratus	ERBB2_HUMAN	PRN	PRN [2]	
3	51	12 2	1257	1 ERBB2_RAT	P06494 rattus norvegicus	ERBB2_HUMAN	PRN	PRN [2]	
4	2 1	230	1 RNF1_VIBCHICK	Q9kt91 vibrio cholerae	ERBB2_HUMAN	PRN	PRN	PRN [2]	
5	9	703	1 EGFR_CHICK	P13387 gallus gallus	ERBB2_HUMAN	PRN	PRN	PRN [2]	
6	9	2 1	1210	1 EGFR_HUMAN	P00533 homo sapiens	ERBB2_HUMAN	PRN	PRN	
7	7	2 1	1210	1 EGFR_MOUSE	Q01279 mus musculus	ERBB2_HUMAN	PRN	PRN	
8	8	1 9	276	1 DICA_RAT	P47727 rattus norvegicus	ERBB2_HUMAN	PRN	PRN	
9	8	1 9	463	1 YV330_MCTU	Q07035 mycobacterium tuberculosis	ERBB2_HUMAN	PRN	PRN	
10	8	1 9	725	1 NCA2_MOUSE	P13594 mus musculus	ERBB2_HUMAN	PRN	PRN	
11	8	8 58	1 NCA1_RAT	P13595 ratmus norvegicus	ERBB2_HUMAN	PRN	PRN	PRN	
12	8	1 9	1115	1 NCAL_MOUSE	P58191 sulfolobus solis	ERBB2_HUMAN	PRN	PRN	
13	7	1 7	62	1 SRCE_SULISO	Q00994 homo sapiens	ERBB2_HUMAN	PRN	PRN	
14	7	1 7	111	1 NADP_HUMAN	Q95336 tupéa gili	ERBB2_HUMAN	PRN	PRN	
15	7	1 7	114	1 GON2_TIFGB	P18941 gallus gallus	ERBB2_HUMAN	PRN	PRN	
16	7	1 7	173	1 NUB6_CHICK	P03038 alcaligenes p48801 galinus gallus	ERBB2_HUMAN	PRN	PRN	
17	7	1 7	192	1 URRB_ACCEU	P09hc801 homosapiens	ERBB2_HUMAN	PRN	PRN	
18	7	1 7	220	1 RGF3_CHICK	P22104 salmonella typhimurium	ERBB2_HUMAN	PRN	PRN	
19	7	1 7	221	1 SDF1_HUMAN	P09hc801 homosapiens	ERBB2_HUMAN	PRN	PRN	
20	7	1 7	224	1 TCTD_SALTY	P09btj7 homo sapiens	ERBB2_HUMAN	PRN	PRN	
21	7	1 7	228	1 YU02_HUMAN	P58320 escherichia coli	ERBB2_HUMAN	PRN	PRN	
22	7	1 7	248	1 DSBB_ECOLI	P77202 escherichia coli	ERBB2_HUMAN	PRN	PRN	
23	7	1 7	248	1 HMGFL_PIVB	P26586 trypanosoma brucei	ERBB2_HUMAN	PRN	PRN	
24	7	1 7	271	1 YCXE_BACSU	P40420 bacillus subtilis	ERBB2_HUMAN	PRN	PRN	
25	7	1 7	287	1 SRPL_SNPN7	Q55032 synnechococcus P46841 mycobacterium	ERBB2_HUMAN	PRN	PRN	
26	7	306	1 MPFL_MCMCLE	P10404 mus musculus	ERBB2_HUMAN	PRN	PRN	PRN	
27	7	307	1 ENV1_MOUSE	C3P1_BACSU	ERBB2_HUMAN	PRN	PRN	PRN	
28	7	313	1 G3P1_BACSU	P09124 bacillus subtilis	ERBB2_HUMAN	PRN	PRN	PRN	
29	7	334	1 XERD_SELRU	P05655 selenomonas prausnitzii	ERBB2_HUMAN	PRN	PRN	PRN	
30	7	341	1 MULTC_HAEIN	P44049 haemophilus prausnitzii	ERBB2_HUMAN	PRN	PRN	PRN	
31	7	357	1 YM27_YEAST	P46955 saccharomyces cerevisiae	ERBB2_HUMAN	PRN	PRN	PRN	
32	7	365	1 IRKA_RAT	P46955 r alpha-sensin	ERBB2_HUMAN	PRN	PRN	PRN	
33	7	379				PRN	PRN	PRN	

-!- PTM: LIGAND-BINDING INCREASES PHOSPHORYLATION ON TYROSINE REST DUEES (BY SIMILARITY).

-!- POLYMORPHISM: THERE ARE POSITIONS 654 AND 655. ALLELE B1 (654-ILE-ILE-655) HAS A FREQUENCY OF 0.206; OF 0.782; ALLELE B2 (654-ILE-VAL-655) HAS A FREQUENCY OF 0.012; ALLELE B3 (654-VAL-VAL-655) HAS A FREQUENCY OF 0.012.

-!- SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL: M11767; AAA35808.1; -.

DR EMBL: M11761; AAA35808.1; JOINED.

DR EMBL: M11762; AAA35808.1; JOINED.

DR EMBL: M11763; AAA35808.1; JOINED.

DR EMBL: M11764; AAA35808.1; JOINED.

DR EMBL: M11765; AAA35808.1; JOINED.

DR EMBL: M11766; AAA35808.1; JOINED.

DR EMBL: M11767; AAA35808.1; JOINED.

DR EMBL: M11768; AAA35808.1; JOINED.

DR EMBL: M11769; AAA35808.1; JOINED.

DR EMBL: M11770; AAA35808.1; JOINED.

DR EMBL: M12036; AAA3598.1; -.

DR EMBL: X03363; CAA27060.1; -.

DR PIR: A25491; A25411; FURIN-LIKE.

DR PIR: A24571; A24571.

DR HSSP: P11362; IFGK.

DR Genew; HGNC:3430; ERBB2.

MIM: 164870; -.

DR InterPro; IPR00494; EGFR\_L\_domain.

DR InterPro; IPR00719; Euk\_pk kinase.

DR InterPro; IPR002174; Furin-like.

DR InterPro; IPR001245; Tyr\_pk kinase.

DR SMART; SMD0261; FU; 3.

DR SMART; SM00219; TYR\_C1.

DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.

DR PROSITE; PS00109; PROTEIN\_TYR; 1.

KW Transmembrane; Glycoprotein; Multigene family; Receptor; Signal;

KW Transferase; Tyrosine-protein kinase; ATP-binding; Phosphorylation; Polymorphism.

FT SIGNAL 1 21 POTENTIAL.

FT CHAIN 22 1255 RECEPTOR PROTEIN-TYROSINE KINASE ERBB-2.

FT DOMAIN 22 652 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 653 675 POTENTIAL.

FT DOMAIN 676 1255 CYTOPLASMIC (POTENTIAL).

FT DOMAIN 720 987 PROTEIN KINASE.

FT NP\_BIND 726 734 ATP (BY SIMILARITY).

FT BINDING 753 753 ATP (BY SIMILARITY).

FT ACT\_SITE 845 845 BY SIMILARITY.

FT DISULFID 195 204 BY SIMILARITY.

FT DISULFID 199 212 BY SIMILARITY.

FT DISULFID 220 227 BY SIMILARITY.

FT DISULFID 224 235 BY SIMILARITY.

FT DISULFID 236 244 BY SIMILARITY.

FT DISULFID 240 252 BY SIMILARITY.

FT DISULFID 255 264 BY SIMILARITY.

FT DISULFID 268 295 BY SIMILARITY.

FT DISULFID 299 311 BY SIMILARITY.

FT DISULFID 315 331 BY SIMILARITY.

FT DISULFID 334 338 BY SIMILARITY.

FT DISULFID 511 520 BY SIMILARITY.

FT DISULFID 515 528 BY SIMILARITY.

FT DISULFID 531 540 BY SIMILARITY.

FT DISULFID 544 560 BY SIMILARITY.

FT DISULFID 563 576 BY SIMILARITY.

FT DISULFID 567 584 BY SIMILARITY.

FT DISULFID 600 623 BY SIMILARITY.

FT DISULFID 625 634 BY SIMILARITY.

FT DISULFID 630 642 BY SIMILARITY.

FT MOD\_RES 1139 1139 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).

FT MOD\_RES 1248 1248 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).

FT CARBOHYD 668 68 N-LINED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 124 124 N-LINED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 187 187 N-LINED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 259 259 N-LINED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 530 530 N-LINED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 571 571 N-LINED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 629 629 N-LINED (GLCNAC. . .) (POTENTIAL).

FT VARIANT 654 654 /FTID=VAR\_004077.

FT VARIANT 655 655 I -> V.

FT CONFLICT 1170 1170 /FTID=VAR\_004078.

FT SEQUENCE 1255 AA; 137909 MN; P -> A (IN REF. 2); 39E9DFDA04DCFF962 CRC64;

Query Match 19.8%; Score 83; DB 1; Length 1255;

Best Local Similarity 100.0%; Pred. No. 6.5e-76;

Matches 83; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 199 CKGSROWGESEEDCOSLRTVCAGGCGARCKGKPLPTDCCHQCAAGCTGPKRSQCLACIHF 258

Db 199 CKGSROWGESEEDCOSLRTVCAGGCGARCKGKPLPTDCCHQCAAGCTGPKRSQCLACIHF 258

QY 259 NHSGICELHCPALVNTDPES 281

Db 259 NHSGICELHCPALVNTDPES 281

RESULT 2

ERBB\_MEASU ID ERBB2\_MEASU STANDARD; PRT; 1254 AA.

AC 060557; 15-DEC-1998 (Rel. 37, Created)

DT 15-DEC-1998 (Rel. 37, Last sequence update)

DR 15-JUN-2002 (Rel. 41, Last annotation update)

DR Receptor protein-tyrosine kinase erbB-2 precursor (EC 2.7.1.112)

DE (P189erbB2) (NEU proto-oncogene) (c-erbB-2).

GN ERBB2 OR NEU.

OS Mesocricetus auratus (Golden hamster).

OC Mammalia; Chordata; Choromata; Veretebrata; Euteleostomi;

OC Mesoicetus; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Crictetinae;

OC Mesoicetus.

OX NCBI\_TaxID=10056;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Nerve;

RA MEDLINE=94193007; PubMed=7908275;

RA Nakamura T., Usuijima T., Ishizaka Y., Nagao M., Arai M., Yamazaki Y., Ishikawa T., "Cloning and activation of the Syrian hamster neu proto-oncogene.", Gene 140:251-255(1994).

RL Gene 140:251-255(1994).

CC -!- FUNCTION: ESSENTIAL COMPONENT OF A NEUREGULIN-RECEPTOR COMPLEX, ALTHOUGH NEUREGULINS DO NOT INTERACT WITH IT ALONE. GP30 IS A POTENTIAL LIGAND FOR THIS RECEPTOR. NOT ACTIVATED BY EGF, TGF-

CC ALPHA AND AMPHIREGULIN (BY SIMILARITY).

CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein tyrosine phosphate.

CC -!- SUBUNIT: HETEROODIMER WITH EACH OF THE OTHER ERBB RECEPTORS (POTENTIAL).

CC -!- SUBCELLULAR LOCATION: Type I membrane protein.

CC -!- PTM: LIGAND-BINDING INCREASES PHOSPHORYLATION ON TYROSINE RESIDUES.

CC -!- SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/>) or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).

Qy	Db
231 LPTDCCHEQCRAAGCTGPKHSDCLACLHFNNSGICBLHCPALVTNTDFFES	2811
231 LPTDCCHEQCRAAGCTGPKHSDCLACLHFNNSGICBLHCPALVTNTDFFES	2811

ID	RESULT	ERB2-RAT	STANDARD:	PRT:	1257 AA.
AC	P0694;				
DT	01-JAN-1988	(Rel. 06, Created)			
DT	15-DEC-1998	(Rel. 37, last sequence update)			
DT	15-JUN-2002	(Rel. 41, last annotation update)			
DE	Barnaghi C.I., Hung M.-C., Weinberg R.A.;	Receptor protein-tyrosine kinase erbB2 precursor (EC 2.7.1.112) (p18erbB2), (Nru protein-tyrosine-oncogene) (C-erbB-2) (Epidermal growth factor receptor-related protein).			
DE	Rattus norvegicus (Rat).	erbB2 (OR NEU).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.				
OC	NCBI-Taxid=10116;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Neuroblastoma;				
RX	MEDLINE=86118662;	PUBMED=33945311;			
RA	Barnaghi C.I., Hung M.-C., Weinberg R.A.;	"The neu oncogene encodes an epidermal growth factor receptor-related protein.";			
RT	Nature 319:226-230(1986).				
RN	[2]				
RP	SEQUENCE OF 852-905 FROM N.A.				
RC	TISSUE=Sciatic nerve;				
RX	MEDLINE=91222860;	PUBMED=2025425;			
RA	Lai C., Lemke G.;	"An extended family of protein-tyrosine kinase genes differentially expressed in the vertebrate nervous system.";			
RT	Neuron 6:691-704(1991).				
RL	RN [3]				
RP	STRUCTURE BY NMR OF 650-668.				
RX	MEDLINE=92155181;	PUBMED=13476763;			
RA	Gullick W.J., Bottomley A.C., Loftis F.J., Doak D.G., Mulvey D., Newman R., Crumpton M.J., Sternberg M.J.E., Campbell I.D.;	"Three dimensional structure of the transmembrane region of the proto-oncogenic and oncogenic forms of the neu protein";			
RT	EMBO J. 11:43-48(1992).	"ALTHOUGH NEUREGULINS DO NOT INTERACT WITH IT ALONE. GP30 IS A POTENTIAL LIGAND FOR THIS RECEPTOR. NOT ACTIVATED BY EGF, TGF- $\alpha$ , ALPHA AND AMPHIREGULIN."			
RL	-1- FUNCTION: ESSENTIAL COMPONENT OF A NEUREGULIN-RECEPTOR COMPLEX.				
CC	-1- SUBUNIT: HETEROODIMER WITH EACH OF THE OTHER ERBB RECEPTORS.				
CC	-1- THE CONSTITUTIVELY ACTIVATED ONCOGENIC VARIANT FORMS A HOMODIMER.				
CC	-1- SUBCELLULAR LOCATION: TYPE I membrane protein.				
CC	-1- PBM: LIGAND-BINDING INCREASES PHOSPHORYLATION ON TYROSINE RESIDUES (BY SIMILARITY).				
CC	-1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein tyrosine phosphate.				
CC	-1- SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to license@isb-sib.ch).				
CC	EMBL: X03362; CAA27059.1; ALT_INIT.				
DR	PIR: A24562; IRTND.				
DR	HSSP: P11362; FGK.				
DR	InterPro: IPR00494; EGFR_L_domain.				
DR	InterPro: IPR00719; Euk_Pkinase.				
DR	InterPro: IPR02174; Furin-like.				
DR	InterPro: IPR0245; Tyr_Pkinase.				
DR	InterPro: IPR004019; YLP_motif.				

DR pfam; PF00069; Pkinase; 1.  
 DR pfam; PF0057; Furin-like; 1.  
 DR pfam; PF1030; Recep\_L\_domain; 2.  
 DR pfam; PF0275; YIP; 2.  
 DR prodom; PD00001; Euk\_pk kinase; 1.  
 DR SMART; SM00261; FU; 3.  
 DR SMART; SM00219; TYRK; 1.  
 DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
 DR PROSITE; PS00111; PROTEIN\_KINASE\_DOM; 1.  
 KW Transmembrane; Glycoprotein; Multigene family; Receptor; Signal;  
 transferase; Tyrosine-protein kinase; ATP-binding; Phosphorylation;  
 proto-oncogene; Disease mutation.  
 KW SIGNAL; 1  
 FT DOMAIN; 22  
 FT DOMAIN; 22  
 FT DOMAIN; 655  
 FT DOMAIN; 678  
 FT DOMAIN; 159  
 FT DOMAIN; 473  
 FT DOMAIN; 722  
 FT NP\_BIND; 728  
 FT BINDING; 755  
 FT ACT\_SITE; 847  
 FT DISULFID; 196  
 FT DISULFID; 200  
 FT DISULFID; 221  
 FT DISULFID; 225  
 FT DISULFID; 245  
 FT DISULFID; 241  
 FT DISULFID; 256  
 FT DISULFID; 269  
 FT DISULFID; 300  
 FT DISULFID; 316  
 FT DISULFID; 335  
 FT DISULFID; 513  
 FT DISULFID; 517  
 FT DISULFID; 533  
 FT DISULFID; 546  
 FT DISULFID; 578  
 FT DISULFID; 569  
 FT DISULFID; 589  
 FT DISULFID; 602  
 FT DISULFID; 625  
 FT DISULFID; 628  
 FT MOD\_RES; 1141  
 FT MOD\_RES; 1250  
 FT CARBOHYD; 188  
 FT CARBOHYD; 188  
 FT CARBOHYD; 260  
 FT CARBOHYD; 532  
 FT CARBOHYD; 573  
 FT CARBOHYD; 631  
 FT VARIANT; 661  
 SQ Sequence 1257 AA; 138831 MW; 6129264583011402 CRC64;

RESULT 4

Query Match 12.2%; Score 51; DB 1; Length 1257;  
 Best Local Similarity 100.0%; Pred. No. 2.2e-43; Matches 51; Conservative 0; Indels 0; Gaps 0;

QY 231 LPTDCCHEQCAAGCTGPKHSOOLACHFNHSOCICELHCPALVYNTDFES 281  
 DB 232 LPTDCCHEQCAAGCTGPKHSOOLACHFNHSOCICELHCPALVYNTDFES 282

DE Election transport complex protein INF.

RNFE\_VIBCH STANDARD; PRT; 230 AA.

AC 09KT91; OC  
 DT 15-JUN-2002 (Rel. 41, Created) OX  
 DT 15-JUN-2002 (Rel. 41, Last sequence update) NCBI\_TaxID=9031; [1]  
 DT 15-JUN-2002 (Rel. 41, Last annotation update) RN  
 DE Sequence from N.A. RP  
 MEDLINE=88261272; PubMed=3260329;

GN RNFE OR VC1012.  
 OS Vibrio cholerae.  
 OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.  
 OX NCBI\_TaxID=6666; [1]

SEQUENCE FROM N.A.

RC STRAIN=EL TOR N16991 / Serotype O1;  
 RX MEDLINE=20406833; PubMed=1095201;  
 RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,  
 RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,  
 RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,  
 RA Elmlundova M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,  
 RA McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,  
 RA Galzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,  
 RA Fraser C.M.;  
 RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio  
 cholerae.",  
 RL Nature 406:477-483 (2000).  
 CC -1- FUNCTION: May be part of a membrane complex involved in electron  
 transport (By similarity).  
 CC -1- SUBUNIT: Composed of at least six subunits: rnfA, rnfB, rnfC,  
 CC rnfD, rnfE and rnfG (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane  
 (potential).  
 CC -1- SIMILARITY: BELONGS TO THE NORDE/RNF/E FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to license@isb-sib.ch).

DR EMBL; AE004183; AAF94173.1; -

DR TIGR; VC1012; -

DR InterPro; IPR03667; Rnf\_Nqr.

DR Pfam; PF02108; Rnf\_Nqr; 1.

KW Electron transport; Transmembrane; Inner membrane; Complete proteome.

FT TRANSMEM; 39  
 FT TRANSMEM; 69  
 FT TRANSMEM; 93  
 FT TRANSMEM; 124  
 FT TRANSMEM; 182  
 SQ SEQUENCE 230 AA; 24710 MW; 262DAD792044D769 CRC64;

RESULT 5

Query Match 2.1%; Score 9; DB 1; Length 230;  
 Best Local Similarity 100.0%; Pred. No. 0.24; Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 LLALLPPGA 20  
 DB 184 LLALLPPGA 192

DE Sequence from N.A. DE  
 EGFR\_ECFR\_CHEM STANDARD; PRT; 703 AA.

AC P13387; DT 01-JAN-1990 (Rel. 13, Created)  
 DT 01-JAN-1990 (Rel. 13, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Epidermal growth factor receptor precursor (EC 2.7.1.112) (CER)  
 DE (Fragment).  
 GN EGFR  
 OS Gallus gallus (Chicken).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
 OX NCBI\_TaxID=9031; [1]  
 RP Sequence from N.A. RN  
 MEDLINE=88261272; PubMed=3260329;

RA	Lax I., Johnson A., Howk R., Sap J., Bellot F., Winkler M., Ullrich A., Vennstrom B., Schlessinger J., Givoli D., "Chicken epidermal growth factor (EGF) receptor: cDNA cloning, expression in mouse cells, and differential binding of EGF and transforming growth factor alpha.";	FT CARBOHYD 613 FT CARBOHYD 633 FT CARBOHYD 648 FT NON TER 703 FT SEQUENCE 703 AA:	CARBOHYD 613 CARBOHYD 633 CARBOHYD 648 NON TER 703 SEQUENCE 703 AA:	613 633 648 703 77427 MW:	N-LINKED (GLCNAC. . .) (POTENTIAL). N-LINKED (GLCNAC. . .) (POTENTIAL). N-LINKED (GLCNAC. . .) (POTENTIAL). N-LINKED (GLCNAC. . .) (POTENTIAL).
RT	Mol. Cell. Biol. 8:1970-1978(1988).				
-!-	FUNCTION: THE EGF RECEPTOR MEDIATES THE BIOLOGICAL SIGNAL OF EGF, AND ALSO OF TGF-ALPHA, AMPHIREGULIN, HEPARIN-BINDING EGF, GP30 AND VACCINIA VIRUS GROWTH FACTOR (BY SIMILARITY).				
-!-	CATALYTIC ACTIVITY: ATP + a protein tyrosine tyrosine phosphate.				
CC	-!- SUBCELLULAR LOCATION: Type I membrane protein.				
CC	-!- MISCELLANEOUS: Binding of EGF to the receptor leads to dimerization, internalization of the EGF-receptor complex, induction of the tyrosine kinase activity, stimulation of cell DNA synthesis, and cell proliferation.				
CC	-!- SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement. (See <a href="http://www.isb-sib.ch/announce">http://www.isb-sib.ch/announce</a> or send an email to license@isb-sib.ch).				
CC	EMBL; M20386; AAA48760_1; -;				
DR	InterPro; IPR00494; EGFR_L_domain.				
DR	InterPro; IPR00719; Euk_pk kinase.				
DR	InterPro; IPR02174; Furin-like.				
DR	InterPro; IPR01245; Tyr_pk kinase.				
DR	Pfam; PF0175; Furin-like; 1.				
DR	Pfam; PF01030; RecPep_L_domain; 2.				
DR	SMART; SM0261; FU_4.				
DR	PROSITE; PS00107; PROTEIN_KINASE_ATP; PARTIAL.				
DR	PROSITE; PS00109; PROTEIN_KINASE_TYR; PARTIAL.				
DR	PROSITE; PS50011; PROTEIN_KINASE_DOM; PARTIAL.				
KW	Transmembrane; Glycoprotein; Receptor; Signal; Tyrosine-protein kinase; ATP-binding; Phosphorylation.				
KW	Tyrosine-protein kinase; ATP-binding; Phosphorylation.				
FT	CHAIN 1 >703 EPIDERMAL GROWTH FACTOR RECEPTOR.				
FT	DOMAIN 31 654 EXTRACELLULAR (POTENTIAL).				
FT	TRANSMEM 655 667 CYTOPLASMIC (POTENTIAL).				
FT	DOMAIN 668 >703 POTENTIAL.				
FT	DISULFID 197 206 BY SIMILARITY.				
FT	DISULFID 201 214 BY SIMILARITY.				
FT	DISULFID 222 230 BY SIMILARITY.				
FT	DISULFID 226 238 BY SIMILARITY.				
FT	DISULFID 239 247 BY SIMILARITY.				
FT	DISULFID 243 255 BY SIMILARITY.				
FT	DISULFID 258 267 BY SIMILARITY.				
FT	DISULFID 271 298 BY SIMILARITY.				
FT	DISULFID 302 314 BY SIMILARITY.				
FT	DISULFID 318 333 BY SIMILARITY.				
FT	DISULFID 336 340 BY SIMILARITY.				
FT	DISULFID 353 352 BY SIMILARITY.				
FT	DISULFID 517 530 BY SIMILARITY.				
FT	DISULFID 533 542 BY SIMILARITY.				
FT	DISULFID 546 562 BY SIMILARITY.				
FT	DISULFID 565 581 BY SIMILARITY.				
FT	DISULFID 589 601 BY SIMILARITY.				
FT	DISULFID 605 627 BY SIMILARITY.				
FT	DISULFID 630 638 BY SIMILARITY.				
FT	DISULFID 634 646 BY SIMILARITY.				
FT	CARBOHYD 134 134 N-LINKED (GLCNAC. . .) (POTENTIAL).				
FT	CARBOHYD 190 190 N-LINKED (GLCNAC. . .) (POTENTIAL).				
FT	CARBOHYD 200 200 N-LINKED (GLCNAC. . .) (POTENTIAL).				
FT	CARBOHYD 359 359 N-LINKED (GLCNAC. . .) (POTENTIAL).				
FT	CARBOHYD 368 368 N-LINKED (GLCNAC. . .) (POTENTIAL).				
FT	CARBOHYD 420 420 N-LINKED (GLCNAC. . .) (POTENTIAL).				
FT	CARBOHYD 573 573 N-LINKED (GLCNAC. . .) (POTENTIAL).				
FT	CARBOHYD 578 578 N-LINKED (GLCNAC. . .) (POTENTIAL).				
RESULT 6					
ID	EGFR_HUMAN				
DR	EGFR_HUMAN STANDARD				
DR	P00533; P02669; Q14225; Q9UMD7; Q9UMG5; Q92795; 000732; P00688; Q92652; Q9R2C9; Q9GXL1; Q9H3C9; 21-JUL-1986 (Rel. 01, Created)				
DT	01-NOV-1997 (Rel. 35, Last sequence update)				
DT	15-JUN-2002 (Rel. 41, Last annotation update)				
DE	Epidermal growth factor receptor precursor (EC 2.7.1.112) (Receptor protein-Tyrosine kinase Erbb-1).				
GN	EGFR OR ERBB1.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
OX	NCB-I-TAXID=9606;				
RN	[1]				
RP	SEQUENCE FROM N.A. (ISOFORM 1).				
RX	MEDLINE:8421929; PubMed=6328312;				
RA	Ullrich A., Coussens L., Havlick J.S., Dull T.J., Gray A., Tam A.W., Lee J., Yarden Y., Libermann T.A., Schlessinger J., Downward J., Mayes E.I.V., Whittle N., Waterfield M.D., Seeburg P.H.;				
RA	"Human epidermal growth factor receptor receptor cdna sequence and aberrant expression of the amplified gene in A431 epidermoid carcinoma cells.";				
RT	Nature 309:418-425(1984).				
RN	[2]				
RP	SEQUENCE FROM N.A. (ISOFORM 2).				
RC	TISSUE=Placenta;				
RX	MEDLINE:9538257; PubMed=7654368;				
RA	Ilekis J.V., Stark B.C., Scoccia B.;				
RA	"Possible role of variant RNA transcripts in the regulation of epidermal growth factor receptor expression in human placenta.";				
RT	Mol. Reprod. Dev. 41:149-156(1995).				
RN	[3]				
RP	SEQUENCE FROM N.A. (ISOFORM 2).				
RC	TISSUE=Placenta;				
RX	MEDLINE:9708086; PubMed=8918811;				
RA	Reiter J.L., Maile N.J.;				
RA	"A 1.8 kb alternative transcript from the human epidermal growth factor receptor gene encodes a truncated form of the receptor-like factor receptor gene.";				
RT	Nucleic Acids Res. 24:4050-4056(1996).				
RN	[4]				
RP	SEQUENCE FROM N.A. (ISOFORM 2).				
RC	TISSUE=Placenta;				
RX	MEDLINE:9725647; PubMed=9103388;				
RA	Ilekis J.V., Gariti J., Niederberger C., Scoccia B.;				
RA	"Expression of a truncated epidermal growth factor receptor-like protein (TEGR) in ovarian cancer.";				
RL	Gynecol. Oncol. 65:36-41(1997).				
RN	[5]				
RP	SEQUENCE FROM N.A. (ISOFORMS 3 AND 4).				
RC	TISSUE=Placenta;				
RX	MEDLINE:21100872; PubMed=11161793;				
RA	Reiter J.L., Threadgill D.W., Eley G.D., Strunk K.E., Danielsen A.J., Schehl Sinclair C., Pearsall R.S., Green P.J., Yee D., Lampland A.L., Balasubramanian S., Crossey T.D., Magnuson T.R., James C.D., Maini N.J.;				
RA	"Comparative genomic sequence analysis and isolation of human and mouse alternative EGFR transcripts encoding truncated receptor				

RT isoforms.";  
 RL Genomics 71:1-20(2001).  
 RN [6]

SEQUENCE OF 575-687 FROM N.A.  
 RT Reiter J.L., Threadgill D.W., Danielsen A.J., Schehl C.M.,  
 RA Lampard A.I., Balasubramanian S., Crossley T.O., Magnusson T.R.,  
 RA Maihle N.J.;  
 "Human and mouse alternative EGFR transcripts encoding only the  
 extracellular domain of the receptor.";  
 Submitted (FEB-1999) to the EMBL/GenBank/DDBJ databases.

SEQUENCE OF 713-924 FROM N.A.  
 RX MEDLINE=84196372; PubMed=6326201;  
 RA Lin C.R., Chen W.S., Krugler W., Stolarsky L.S., Weber W.,  
 Evans R.M., Verma I.M., Gill G.N., Rosenthal M.G.;  
 "Expression cloning of human EGF receptor complementary DNA: gene  
 amplification and three related messenger RNA products in A431  
 cells.";  
 Science 224:843-848(1984).  
 RN [8]

SEQUENCE OF 150-962 FROM N.A.  
 RP SEQUENCE OF 150-962 FROM N.A.  
 RX MEDLINE=84245835; PubMed=6330563;  
 RA Xu Y.-H., Ishii S., Clark A.J.L., Sullivan M., Wilson R.K., Ma D.P.,  
 Roe B.A., Merlino G.T., Pastan I.;  
 "Human epidermal growth factor receptor receptor cDNA is homologous to a  
 variety of RNAs overproduced in A431 carcinoma cells.";  
 Nature 309:806-810(1984).  
 RN [9]

SEQUENCE OF 1028-1210 FROM N.A.  
 RX MEDLINE=85046483; PubMed=093780;  
 RA Simmen F.A., Gope M.L., Schulz T.Z., Wright D.A., Carpenter G.,  
 O'Malley B.W.;  
 "Isolation of an evolutionarily conserved epidermal growth factor  
 receptor cDNA from human A431 carcinoma cells";  
 RL Biochem. Biophys. Res. Commun. 124:125-132(1984).  
 RN [10]

SEQUENCE OF 1-29 FROM N.A.  
 RP MEDLINE=88217333; PubMed=3329716;  
 RX Haley J.D., Whittle N., Bennett P., Kinchington D., Ulrich A.,  
 Waterfield M.D.;  
 "The human EGF receptor gene: structure of the 110 kb locus and  
 identification of sequences regulating its transcription.";  
 RL Oncogene Res. 1:375-396(1987).  
 RN [11]

SEQUENCE OF 1-29 FROM N.A.  
 RX MEDLINE=91107677; PubMed=1988448;  
 RA Haley J.D., Waterfield M.D.;  
 "Contributory effects of de novo transcription and premature  
 transcript termination in the regulation of human epidermal growth  
 factor receptor proto-oncogene RNA synthesis.";  
 RL J. Biol. Chem. 266:1746-1753(1991).  
 RN [12]

SEQUENCE OF 1-29 FROM N.A.  
 RX MEDLINE=85270438; PubMed=2991899;  
 RA Ishii S., Xu Y.H., Stratton R.H., Roe B.A., Merlino G.T., Pastan I.;  
 "Characterization and sequence of the promoter region of the human  
 epidermal growth factor receptor gene.";  
 Proc. Natl. Acad. Sci. U.S.A. 82:4920-4924(1985).  
 RN [13]

SEQUENCE OF 540.  
 RA Kohda D.;  
 "Submitted (SEP-1997) to the SWISS-PROT data bank.".

RT RECEPTOR ACTIVITY.  
 RX MEDLINE=84191554; PubMed=63225948;  
 RA Mroczkowski B., Mosig G., Cohen S.;  
 "ATP stimulated interaction between epidermal growth factor receptor  
 and supercoiled DNA.";  
 Nature 309:270-273(1984).  
 RN [15]

PHOSPHORYLATION.  
 RX MEDLINE=89278137; PubMed=2543678;  
 RA Margolis B.L., Lax I., Kris R., Dombalagian M., Honegger A.M.,

RA Hawk R., Girola D., Ulrich A., Schlessinger J.;  
 RT "All autophasphorylation sites of epidermal growth factor (EGF)  
 receptor and HER2/neu are located in their carboxyl-terminal tails.  
 RT Identification of a novel site in EGFR receptor.";  
 RL J. Biol. Chem. 264:10667-10671(1989).  
 RN [16]

RP CARBOHYDRATE-LINKAGE SITES ASN-128; ASN-175; ASN-413; ASN-444 AND  
 ASN-528.  
 RX MEDLINE=96398132; PubMed=8962717;  
 RA Smith K.D., Davies M.J., Bailey D., Renouf D.V., Hounsell E.F.;  
 RT "Analysis of the glycosylation patterns of the extracellular domain of  
 the epidermal growth factor receptor expressed in Chinese hamster  
 ovary fibroblasts.";  
 RL Growth Factors 13:121-132(1996).  
 RN [17]

RP CARBOHYDRATE-LINKAGE SITES ASN-56; ASN-352; ASN-361; ASN-568 AND  
 ASN-603.  
 RX MEDLINE=20198209; PubMed=10731668;  
 RA Sato C., Kim J.-H., Abe Y., Saito K., Yokoyama S., Kohda D.;  
 RT "Characterization of the N-oligosaccharides attached to the atypical  
 Asn-X-Cys sequencing of recombinant human epidermal growth factor  
 receptor.";  
 RL J. Biochem. 127:65-72(2000).  
 RN [18]

RP PARTIAL SEQUENCE, AND DISULFIDE BONDS.  
 RX MEDLINE=98225195; PubMed=9556602;  
 RA Abe Y., Odaka M., Inagaki F., Lax I., Schlessinger J., Kohda D.;  
 RT "Disulfide bond structure of human epidermal growth factor receptor.";  
 RL J. Biol. Chem. 273:11150-11157(1998).  
 RN [19]

RP REVIEW.  
 RX MEDLINE=87297456; PubMed=3039909;  
 RA Carpenter G.;  
 RT "Receptors for epidermal growth factor and other polypeptide  
 mitogens.";  
 RL Annu. Rev. Biochem. 56:881-914(1987).

CC -I- FUNCTION: Receptor for EGF, but also for other members of the EGF  
 family, as TGF-alpha, amphiregulin, betacellulin, heparin-binding  
 EGF-like growth factor, Gp30 and vaccinia virus growth factor. Is  
 involved in the control of cell growth and differentiation.

CC -I- FUNCTION: Isoform 2/truncated isoform may act as an antagonist.

CC -I- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein  
 tyrosine phosphate.

CC -I- SUBCELLULAR LOCATION: Type 1 membrane protein. Isoform 2 is  
 secreted.

CC -I- ALTERNATIVE PRODUCTS: 4 isoforms; 1/p170 (shown here), 2/p60/  
 truncated isoform/EGFR, 3/p110 and 4; are produced by  
 alternative splicing.

CC -I- TISSUE SPECIFICITY: Expressed in placenta. Isoform 2 is also  
 expressed in ovarian cancers.

CC -I- MISCELLANEOUS: Binding of EGF to the receptor leads to  
 dimerization, internalization of the EGF-receptor complex,  
 induction of the tyrosine kinase activity, stimulation of cell DNA  
 synthesis, and cell proliferation.

CC -I- SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to license@isb-sib.ch).

CC DR EMBL: X05088; CAA25240.1; -  
 DR EMBL: U95098; AAB53053.1; -  
 DR EMBL: U48722; AAC50802.1; -  
 DR EMBL: U48723; AAC50804.1; -  
 DR EMBL: U48724; AAC50796.1; -  
 DR EMBL: U48725; AAC50797.1; -  
 DR EMBL: U48726; AAC50798.1; -





Db	315 PLQRIV 322	DR	InterPro; IPR003006; Ig_MHC.
RESULT	10	DR	InterPro; IPR003598; Ig_c2.
NCA2_MOUSE	STANDARD;	DR	Pfam; PF00041; fn3; 2.
ID	NCA2_MOUSE	DR	Pfam; PF00047; Ig; 5.
AC	P13594; 061950; 01-JAN-1990 (Rel. 13, Last sequence update)	DR	SMART; SM00050; FN3; 2.
DT	15-JUN-2002 (Rel. 41, Last annotation update)	DE	SMART; SM00408; IgC2; 5.
DE	Neural cell adhesion molecule 1, 120 kDa isoform precursor (N-CAM 120)	DE	(NCAM-120).
GN	NCAM OR NCAM.	KW	Cell adhesion; Glycoprotein; Repeat; Alternative splicing; Immunoglobulin domain; Signal; Heparin binding; GPI-anchor.
OS	Mus musculus (Mouse).	KW	IMMUNOGLOBULIN DOMAIN; SIGNAL.
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	FT	NEURAL CELL ADHESION MOLECULE 1, 120 KDA ISOFORM.
OX	NCBI_TAXID=10090;	FT	IG-LIKE C2-TYPE DOMAIN 1.
RN	[1]	FT	IG-LIKE C2-TYPE DOMAIN 2.
RP	SEQUENCE FROM N.A.	FT	IG-LIKE C2-TYPE DOMAIN 3.
RX	STRAIN=C57BL/6;	FT	IG-LIKE C2-TYPE DOMAIN 4.
RX	MEDLINE=87246524; PubMed=3595563;	FT	IG-LIKE C2-TYPE DOMAIN 5.
RA	Barthels D., Santoni M.-J., Wille W., Ruppert C., Caix J.-C.,	FT	FIBRONECTIN TYPE-III 1.
RA	Hirsch M.-R., Fontecilla-Camps J.-C., Goridis C.,	FT	FIBRONECTIN TYPE-III 2.
RT	"Isolation and nucleotide sequence of mouse NCAM cDNA that codes for a Mr 79,000 polypeptide without a membrane-spanning region.";	FT	HEPARIN-BINDING (POTENTIAL).
RL	EMBO J. 6:907-914(1987).	FT	HEPARIN-BINDING (POTENTIAL).
RN	[2]	FT	PROBABLE.
RP	SEQUENCE OF 20-700 FROM N.A.	FT	PROBABLE.
RC	STRAIN=C57BL/6; TISSUE=brain;	FT	PROBABLE.
RX	MEDLINE=89251563; PubMed=27212486;	FT	PROBABLE.
RA	Santoni M.J., Barthels D., Wille M., Goridis C., Wille M.;	FT	PROBABLE.
RT	"differential exon usage involving an unusual splicing mechanism generates at least eight types of NCAM cDNA in mouse brain.";	FT	PROBABLE.
RL	EMBO J. 8:385-392(1989).	FT	PROBABLE.
RN	[3]	FT	PROBABLE.
RP	SEQUENCE OF 612-725 FROM N.A.	FT	PROBABLE.
RX	MEDLINE=88285628; PubMed=3306534;	FT	PROBABLE.
RA	Barbas J.A., Chaik J.C., Steinmetz M., Goridis C.;	FT	PROBABLE.
RT	"Differential splicing and alternative polyadenylation generates distinct NCAM transcripts and proteins in the mouse.";	FT	PROBABLE.
RL	EMBO J. 7:625-632(1988).	FT	PROBABLE.
RN	[4]	FT	PROBABLE.
RP	SEQUENCE OF 20-36.	FT	PROBABLE.
RX	MEDLINE=86140120; PubMed=3552556;	FT	PROBABLE.
RA	Rougon G., Marshak D.R.;	FT	PROBABLE.
RT	"Structural and immunological characterization of the amino-terminal domain of mammalian neural cell adhesion molecules.";	FT	PROBABLE.
RL	J. Biol. Chem. 261:3396-3401(1986).	FT	PROBABLE.
CC	-1- FUNCTION: THIS PROTEIN IS A CELL ADHESION MOLECULE INVOLVED IN NEURON-NEURON ADHESION, NEURITE FASCICULATION, OUTGROWTH OF NEURITES, ETC.	FT	PROBABLE.
CC	-1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.	FT	PROBABLE.
CC	-1- ALTERNATIVE PRODUCTS: 3 isoforms; N-CAM 180 (AC P13595), N-CAM 140 (AC P13595), and N-CAM 120 (shown here); are produced by alternative splicing.	FT	PROBABLE.
CC	-1- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY.	FT	PROBABLE.
CC	-1- SIMILARITY: CONTAINS 5 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.	FT	PROBABLE.
CC	-1- SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE III-LIKE DOMAINS.	FT	PROBABLE.
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to license@lsb-sib.ch).	FT	PROBABLE.
CC	SEQUENCE FROM N.A.	FT	PROBABLE.
DR	EMBL; Y00051; CAA60263.1; -	FT	PROBABLE.
DR	EMBL; X15049; CAA33148.1; ALT_SEQ.	FT	PROBABLE.
DR	EMBL; X01795; CAA30173.1; -	FT	PROBABLE.
DR	PIR; A29673; IJMSN.	FT	PROBABLE.
DR	MGD; MG-197281; NCAML.	FT	PROBABLE.
DR	InterPro; IPR003961; FN_LIII.	FT	PROBABLE.
RESULT	11	DR	InterPro; IPR003006; Ig_MHC.
NCAL_RAT	STANDARD;	DR	InterPro; IPR003598; Ig_c2.
ID	NCAL_RAT	DR	Pfam; PF00041; fn3; 2.
AC	P13596;	DR	Pfam; PF00047; Ig; 5.
DT	01-JAN-1990 (Rel. 13, Last sequence update)	DR	SMART; SM00050; FN3; 2.
DT	15-JUN-2002 (Rel. 41, Last annotation update)	DE	SMART; SM00408; IgC2; 5.
DE	Neural cell adhesion molecule 1, 140 kDa isoform precursor (N-CAM 140)	DE	(NCAM-140).
GN	NCAM OR NCAM.	DE	NCAM OR NCAM.
OS	Rattus norvegicus (Rat).	DE	NCAM OR NCAM.
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.	DE	NCAM OR NCAM.
OX	NCBI_TAXID=10116;	DE	NCAM OR NCAM.
RN	[1]	DE	NCAM OR NCAM.
RP	SEQUENCE FROM N.A.	FT	PROBABLE.
RC	TISSUE=Brain;	FT	PROBABLE.
RX	MEDLINE=8805925; PubMed=3680385;	FT	PROBABLE.
RA	Small S.J., Shull G.E., Santoni M.-J., Akeson R.;	FT	PROBABLE.
RT	"Identification of a cDNA clone that contains the complete coding sequence for a 140-kDa rat NCAM polypeptide.";	FT	PROBABLE.
J. Cell Biol. 105:2335-2345(1987).	FT	PROBABLE.	

[1] RN SEQUENCE OF 355-364 FROM N.A.  
 RP NCAM\_MOUSE STANDARD; PRT; 1115 AA.  
 RX MEDLINE=90166485; PubMed=2483093;  
 RA Small S.J., Haines S.L., Akeson R.A.;  
 RT "Polyepitope variation in an N-CAM extracellular immunoglobulin-like  
 fold is developmentally regulated through alternative splicing.";  
 RL Neuron 1:1007-1017(1988).  
 CC FUNCTION: THIS PROTEIN IS A CELL ADHESION MOLECULE INVOLVED IN  
 NEURON-NEURON ADHESION, NEURITE FASCICULATION, OUTGROWTH OF  
 NEURITES, ETC.  
 CC -I- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -I- ALTERNATIVE PRODUCTS: A number of isoforms are produced by  
 alternative splicing.  
 CC -I- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY.  
 CC -I- SIMILARITY: CONTAINS 5 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.  
 CC -I- SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE III-LIKE DOMAINS.

[2] RN This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement see <http://www.isb-sib.ch/announce/>  
 or send an email to license@isb-sib.ch).

CC EMBL; X06564; CA2A9809; 1; -.  
 DR EMBL; M32611; AAA41679; 1; -.  
 DR PRK; S00846; IJRNIC.  
 DR InterPro; IPR003961; FN\_III.  
 DR InterPro; IPR03006; IG\_MHC.  
 DR InterPro; IPR003598; Ig\_C2.  
 DR Pfam; PF00041; fn3; 2.  
 DR Pfam; PF00047; ig; 5.  
 DR SMART; SM00060; FN3; 2.  
 DR SMART; SM00408; IgC2; 5.  
 KW Cell adhesion; Glycoprotein; Transmembrane; Repeat;  
 KW Immunoglobulin domain; Alternative splicing; Signal; Heparin-binding.  
 FT SIGNAL 1 19  
 FT CHAIN 20 858 NEURAL CELL ADHESION MOLECULE 1, 140 KDA  
 FT ISOFORM.  
 FT DOMAIN 20 721 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 722 739 POTENTIAL.  
 FT DOMAIN 740 858 CITOPLASMIC (POTENTIAL).  
 FT DOMAIN 740 858 IG-LIKE C2-TYPE DOMAIN 1.  
 FT DOMAIN 132 103 IG-LIKE C2-TYPE DOMAIN 2.  
 FT DOMAIN 228 295 IG-LIKE C2-TYPE DOMAIN 3.  
 FT DOMAIN 323 403 IG-LIKE C2-TYPE DOMAIN 4.  
 FT DOMAIN 430 497 IG-LIKE C2-TYPE DOMAIN 5.  
 FT DOMAIN 514 615 FIBRONECTIN TYPE-III 1.  
 FT DOMAIN 616 712 FIBRONECTIN TYPE-III 2.  
 FT DOMAIN 152 156 HEPARTIN-BINDING (POTENTIAL).  
 FT DOMAIN 161 165 HEPARIN BINDING (POTENTIAL).  
 FT DISULFID 41 96 BY SIMILARITY.  
 FT DISULFID 139 189 BY SIMILARITY.  
 FT DISULFID 235 288 BY SIMILARITY.  
 FT DISULFID 330 396 BY SIMILARITY.  
 FT DISULFID 437 490 BY SIMILARITY.  
 FT CARBOHYD 222 222 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 316 316 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 348 348 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 434 434 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 460 460 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 489 489 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT SEQUENCE 858 AA; 94658 MW; ERA106A4EA0550F6 CRC64;

Query Match Score 8; DB 1; Length 858;  
 Best Local Similarity 100.0%; Pred. No. 7.8; O; Indels 0; Gaps 0;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 322 VTAEDGTO 329  
 |||||  
 Db 98 VTAEDGTO 105

RESULT 12  
 NCAL\_MOUSE STANDARD; PRT; 1115 AA.  
 ID NCAL\_MOUSE STANDARD; PRT; 1115 AA.  
 AC P13955; Q61049; 13, Created  
 DT 01-JUL-1993 (Rel. 26, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Neural cell adhesion molecule 1, 180 kDa isoform precursor (N-CAM 180)  
 DE (NCAM-180).  
 GN NCAM OR NCAM.  
 OS MUS musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OC NCBI\_TAXID=10090;  
 RN [1] SEQUENCE FROM N.A. (ISOFORM N-CAM 180).  
 RP STRAIN=C57BL/6;  
 RX MEDLINE=87246524; PubMed=3595563;  
 RA Barthels D., Santoni M.J., Wille W., Ruppert C., Chaix J.C.,  
 RA Hirsch M.R., Fontecilla-Camps J.C., Goridis C.;  
 RA "Isolation and nucleotide sequence of mouse NCAM cDNA that codes for  
 a Mr 79,000 polypeptide without a membrane-spanning region."  
 RT EMBO J. 6: 907-914 (1987).  
 RL [2] SEQUENCE OF 529-1115 FROM N.A. (ISOFORM N-CAM 140).  
 RP STRAIN=C57BL/6;  
 RX MEDLINE=88067687; PubMed=3684567;  
 RA Santoni M.-J., Barthels D., Barbas J.A., Hirsch M.-R., Steinmetz M.,  
 RA Goridis C., Wille W.;  
 RA "Analysis of cDNA clones that code for the transmembrane forms of the  
 mouse neural cell adhesion molecule (NCAM) and are generated by  
 alternative RNA splicing.";  
 RT Nucleic Acids Res. 15:8621-8641(1987).  
 RL [3] SEQUENCE OF 642-1115 FROM N.A. (ISOFORM N-CAM 180).  
 RP MEDLINE=88233628; PubMed=3396534;  
 RA Barbas J.A., Chaix J.C., Steinmetz M., Goridis C.;  
 RT "Differential splicing and alternative polyadenylation generates  
 distinct NCAM transcripts and proteins in the mouse.";  
 RL EMBO J. 7:625-632(1988).  
 RN [4] SEQUENCE OF 804-1081 FROM N.A. (ISOFORM N-CAM 180).  
 RP STRAIN=C57BL/6;  
 RX MEDLINE=242737; PubMed=2454455;  
 RA Barthels D., Vopper G., Wille W.;  
 RT "NCAM-180, the large isoform of the neural cell adhesion molecule of  
 the mouse, is encoded by an alternatively spliced transcript.";  
 RL Nucleic Acids Res. 16:4217-4225(1988).  
 RN [5] SEQUENCE OF 702-1115 FROM N.A. (ISOFORM N-CAM 140).  
 RP STRAIN=C57BL/6; TISSUE=Brain;  
 RX MEDLINE=89251563; PubMed=2721486;  
 RA Santoni M.J., Barthels D., Vopper G., Boned A., Goridis C., Wille M.;  
 RT "The differential exon usage involving an unusual splicing mechanism  
 generates at least eight types of NCAM cDNA in mouse brain.";  
 RL EMBO J. 8:385-392(1989).  
 RN [6] SEQUENCE OF 20-36.  
 RP MEDLINE=86140120; PubMed=3512556;  
 RX MEDLINE=86140120; PubMed=3512556;  
 RA Rougon G., Marshak D.R.;  
 RT "Structural and immunological characterization of the amino-terminal  
 domain of mammalian neural cell adhesion molecules.";  
 RL J. Biol. Chem. 261:3396-3401(1986).  
 CC -I- FUNCTION: THIS PROTEIN IS A CELL ADHESION MOLECULE INVOLVED IN  
 CC NEURON-NEURON ADHESION, NEURITE FASCICULATION, OUTGROWTH OF  
 CC NEURITES, ETC.  
 CC -I- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -I- ALTERNATIVE PRODUCTS: 3 isoforms; N-CAM 180 (shown here), N-CAM  
 CC 140 and N-CAM 120 (AC P13594); are produced by alternative  
 CC splicing.  
 CC -I- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY.  
 CC -I- SIMILARITY: CONTAINS 5 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.

-!- SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE III-LIKE DOMAINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@ib-sib.ch).

CC

DR EMBL; X07200; CAA30177.1; -

DR EMBL; Y00051; -; INON\_ANNOTATED\_CDS.

DR EMBL; X05328; CAA29641.1; -

DR EMBL; X07195; CAA0173.1; -

DR EMBL; X07244; CAA30230.1; -

DR EMBL; X15051; CAA33150.1; -

DR PIR; A29673; IJMSNL.

DR MGD; MGI:97281; NCaml.

DR InterPro; IPR00361; FN\_III.

DR InterPro; IPR00306; Ig\_MHC.

DR InterPro; IPR003598; Ig\_c2.

DR Pfam; PF00041; fn3; 2.

DR Pfam; PF00047; ig; 5.

DR SMART; SM00060; FN3; 2.

DR SMART; SM00408; IgC2; 5.

KW Cell adhesion; Glycoprotein; Transmembrane; Repeat;

KW Immunoglobulin domain; Alternative splicing; Signal; Heparin-binding.

FT SIGNAL 1 19

FT CHAIN 20 1115

NEURAL CELL ADHESION MOLECULE 1, 180 kDa

FT DOMAIN 20 711

FT TRANSMEM 712 729

FT DOMAIN 730 1115

FT DOMAIN 34 103

FT DOMAIN 132 196

FT DOMAIN 228 295

FT DOMAIN 323 393

FT DOMAIN 420 487

FT DOMAIN 519 596

FT DOMAIN 625 692

FT DOMAIN 152 156

FT DOMAIN 161 165

FT DISULFID 41 96

FT DISULFID 139 189

FT DISULFID 235 288

FT DISULFID 330 386

FT DISULFID 427 480

FT CARBOHYD 222 222

FT CARBOHYD 316 316

FT CARBOHYD 348 348

FT CARBOHYD 424 424

FT CARBOHYD 450 450

FT CARBOHYD 479 479

FT VARIOPLIC 810 1076

SQ SEQUENCE 1115 AA; 119351 MW; 2C93DCD474CFBCAF CRC64;

Query Match 1..9%; Score 8; DB 1; Length 1115;

Best Local Similarity 100.0%; Pred. No. 9.8;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 322 VTAEDGTQ 329

Db 98 VTAEDGTQ 105

RESULT 13

SECE\_SULSO STANDARD; PRT; 62 AA.

ID SECE\_SULSO STANDARD; PRT; 62 AA.

AC PS8191; 16-OCT-2001 (Rel. 40, Created)

AC DT 16-OCT-2001 (Rel. 40, Last sequence update)

AC DT 16-OCT-2001 (Rel. 40, Last annotation update)

CC DE preprotein translocase sece subunit (Protein transport protein SEC61).

CC DE gamma subunit homolog).

CC GN SECE OR SS05663.

CC OS Sulfolobus solfataricus.

CC OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;

CC OC Sulfolobus.

CC OC NCBI\_TaxID=2287;

CC RN [1] SEQUENCE FROM N.A.

CC RP STRAIN=ATCC 35992 / DSM 1617 / P2;

CC RX MEDLINE=21332296; PubMed=11427726;

CC RA Awayez M.J., Chan-Welsher C.C.-Y., Clausen I.G., Curtis B.A., De Moors A., Erauso G., Fletcher C., Gordon P.M.K., Heikamp-de Jong J., Jeffries A.C., Koza C.J., Medina N., Peng X., Thi-Ngoc H.P., Redder P., Schenk M.E., Theriault C., Tolstrup N., Charlebois R.L., Doolittle W.F., Duquet M., Gaasterland T., Garrett R.A., Ragan C.W., Sensen C.W., Van der Oost J.; \*The complete genome of the crenarchaeon Sulfolobus solfataricus P2.\*; Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840(2001).

CC RL TIGRFAMS; TIGR00327; sece\_euk\_arch.

CC CC -!- FUNCTION: Involved in Protein export (By similarity).

CC CC -!- SUBUNIT: Component of the protein translocase complex (By similarity).

CC CC -!- SUBCELLULAR LOCATION: Tail-anchored membrane protein (By similarity).

CC CC -!- SIMILARITY: BELONGS TO THE SECE/SEC61-GAMMA FAMILY.

CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@ib-sib.ch).

CC DR EMBL; AE006668; AAK0677.1; -

CC DR InterPro; IPR001901; SECE.

CC DR InterPro; IPR004795; SECE\_euk\_arch.

CC DR PROSITE; PS01067; SECE\_SEC61G; FALSE NEG.

CC DR KW Protein transport; Translocation; Transmembrane; Complete proteome.

CC FT TRANSMEM 40 60 POTENTIAL.

CC SO SEQUENCE 62 AA; 7184 MW; CD51874F424BE9FA CRC64;

Query Match 1..7%; Score 7; DB 1; Length 62;

Best Local Similarity 100.0%; Pred. No. 8.2;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 178 QLAFLT 184

Db 54 QLAFLI 60

RESULT 14

NADE\_HUMAN STANDARD; PRT; 111 AA.

AC 00094; 01-JUN-1994 (Rel. 29, Created)

AC DT 01-JUN-1994 (Rel. 29, Last sequence update)

AC DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE P75NTR-associated cell death executor (Nerve growth factor receptor associated protein 1) (Ovarian granulosa cell 13.0 kDa protein HGR74).

DE NGFRAP1 OR NAD.

OS Homo sapiens

OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Buteraria; Primates; Catarrhini; Hominoidea; Homo.

OC OX NCBI\_TaxID=9606;

OC RN [1] SEQUENCE FROM N.A.

RC TISSUE=Ovary.

RC MEDLINE=9102550; PubMed=2171551;

RC RA Rapaport G., Freudenstein J., Klaudiny J., Mucha J., Wempe F., Zimmer M., Scheit K.H.;

RT "Characterization of three abundant mRNAs from human ovarian granulosa cells."  
 RL DNA Cell Biol. 9:479-485(1990).  
 RN [2]  
 RP SEQUENCE FROM N.A., AND CHARACTERIZATION.  
 RX MEDLINE=2028829; Published=1076472;  
 RA Mukai J., Hachiya T., Shoji-Hoshino S., Kimura M.T., Nadano D.,  
 RA Suvanto P., Hanacka T., Li Y., Irie S., Greene D.A., Sato T.A.;  
 RT "NADe, a p75NTR-associated cell death executor, is involved in signal transduction mediated by the common neurotrophin receptor p75NTR."  
 RL J. Biol. Chem. 275:17566-17570(2000).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Eye;  
 RA Strausberg R.;  
 RL Submitted (FEB-2001) to the EMBL/GenBank/DDBJ databases.  
 CC -!- FUNCTION: May be a signaling adaptor molecule involved in p75NTR-mediated apoptosis induced by NGF. May play an important role in the pathogenesis of neurogenetic diseases.  
 CC -!- SUBUNIT: Binds to the DEATH domain of p75NTR/NGFR.  
 CC -!- TISSUE SPECIFICITY: FOUND IN OVARIAN GRANULOSA CELLS, TESTIS, PROSTATE AND SEMINAL VESICLE TISSUE.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC DR U63327; AAB16838\_1;  
 CC DR InterPro; IPR001202; GnrH.  
 CC DR Pfam; PF00446; GnrH; 1.  
 CC DR PROSITE; PS00473; GNRH; 1.  
 CC KW Cleavage on pair of basic residues; Hormone; Amidation; Signal.  
 FT SIGNAL 1 25 BY SIMILARITY.  
 FT CHAIN 26 114 PROGNADOLIBERIN II.  
 FT PEPTIDE 26 35 GONADOLIBERIN II.  
 FT PEPTIDE 39 114 GNRH-ASSOCIATED PEPTIDE II.  
 FT MOD\_RES 26 26 PYRROLIDONE CARBOXYLIC ACID  
 FT MOD\_RES 35 35 (BY SIMILARITY). AMIDATION (G-36). PROVIDE AMIDE GROUP.  
 SQ SEQUENCE 114 AA: 12123 MW: 680E901C6869EC1 CRC64;  
 DR EMBL; M38188; AA063232\_1; -.  
 DR EMBL; AF187064; AAFF75129\_1; -.  
 DR EMBL; BC003190; AAH03190\_1; -.  
 DR PIR; C35826; C35826.  
 DR Genew; HGNC:11338B; NGFRAP1.  
 DR MIM: 300361.  
 DR InterPro; IPR001230; Prenyl-site.  
 KW Apoptosis.  
 SQ SEQUENCE 111 AA; 12958 MW; 29AA0573282C933E CRC64;

Query Match 1.7%; Score 7; DB 1; length 111;  
 Best Local Similarity 100.0%; Pred. No. 14;  
 Matches 7; Conservative 0; Mismatches 0;  
 Indels 0; Gaps 0;

Qy 137 LREQLQR 143  
 Db 81 LREQLQR 87

RESULT 15  
 GON2\_TUPGB STANDARD; PRT; 114 AA.  
 ID GON2\_TUPGB  
 AC Q95336;  
 DT 15-DEC-1998 (Rel. 37, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE Progonadotropin II precursor [contains: Gonadotropin II (LHRH II) (Luteinizing hormone releasing hormone II); Gonadotropin releasing hormone II) (GnRH II); GnRH-associated peptide III].  
 DE hormone II) (GnRH II) (Lutiberin II).  
 GN GNRH2.  
 OS *tupaia glis belangeri* (Common tree shrew).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Scandentia; Tupaiidae; Tupaiidae; Tupaiidae; NCBI TaxID=9396;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Hypothalamus;  
 RX MEDLINE=97079639; Published=8921350;  
 RA Kasten T.L., White S.A., Norton T.T., Bond C.T., Adelman J.P., Fernald R.D.;  
 RT "Characterization of two new preprognathic mRNAs in the tree shrew: first direct evidence for mesencephalic Gnrh gene expression in a

RT placental mammal." Gen. Comp. Endocrinol. 104:7-19(1996).  
 RL THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING HORMONES.  
 CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- TISSUE SPECIFICITY: MIDRAIN.  
 CC -!- SIMILARITY: BELONGS TO THE GNRH FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - European Bioinformatics Institute. There are no restrictions on its use by non profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).  
 CC DR EMBL; U63327; AAB16838\_1;  
 CC DR InterPro; IPR001202; GnrH.  
 CC DR Pfam; PF00446; GnrH; 1.  
 CC DR PROSITE; PS00473; GNRH; 1.  
 CC KW Cleavage on pair of basic residues; Hormone; Amidation; Signal.  
 FT SIGNAL 1 25 BY SIMILARITY.  
 FT CHAIN 26 114 PROGNADOLIBERIN II.  
 FT PEPTIDE 26 35 GONADOLIBERIN II.  
 FT PEPTIDE 39 114 GNRH-ASSOCIATED PEPTIDE II.  
 FT MOD\_RES 26 26 PYRROLIDONE CARBOXYLIC ACID  
 FT MOD\_RES 35 35 (BY SIMILARITY). AMIDATION (G-36). PROVIDE AMIDE GROUP.  
 SQ SEQUENCE 114 AA: 12123 MW: 680E901C6869EC1 CRC64;  
 DR EMBL; M38188; AA063232\_1; -.  
 DR EMBL; AF187064; AAFF75129\_1; -.  
 DR EMBL; BC003190; AAH03190\_1; -.  
 DR PIR; C35826; C35826.  
 DR Genew; HGNC:11338B; NGFRAP1.  
 DR MIM: 300361.  
 DR InterPro; IPR001230; Prenyl-site.  
 KW Apoptosis.  
 SQ SEQUENCE 111 AA; 12958 MW; 29AA0573282C933E CRC64;

Query Match 1.7%; Score 7; DB 1; length 114;  
 Best Local Similarity 100.0%; Pred. No. 14;  
 Matches 7; Conservative 0; Mismatches 0;  
 Indels 0; Gaps 0;

Qy 347 PRPAAVP 353  
 Db 105 PRPAAVP 111

Search completed: January 14, 2003, 17:13:24  
 Job time : 18.3032 secs

Copyright (c) 1993 - 2003 Compugen Ltd.	GenCore version 5.1.3				
OM protein - protein search, using sw model					
Run on:	January 14, 2003, 17:11:37 ; Search time 24.3996 Seconds (without alignments) 1650.860 Million cell updates/sec				
Title:	US-09-506-079E-2				
Perfect score:	419				
Sequence:	1 MELANICRWGILLALLLPPGA, . . . . . VGRGXDPDAHVAVXLSRYEG 419				
Scoring table:	Oligo Gapop 60.0 , Gapext 60.0				
Searched:	283224 seqs, 96134422 residues				
Word size :	0				
Total number of hits satisfying chosen parameters:	28324				
Minimum DB seq length:	0				
Maximum DB seq length:	200000000				
Post-processing: Listing first 45 summaries					
Database :	PIR_73;*				
1:	PIR1;*				
2:	PIR2;*				
3:	PIR3;*				
4:	PIR4;*				
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.					
SUMMARIES					
Result No.	Score	Query	Match Length	DB ID	Description
1	83	19.8	1255	1	A24571 protein-tyrosine kinase (EC 2.7.1.112) erbB2 precursor - human p-erbB precursor -
2	51	12.2	1254	2	A148161 protein-tyrosine kinase receptor with extensive homology to EGF receptor shares chro
3	51	12.2	1260	1	TVTRNU Rnf5-related prote
4	9	2.1	230	2	B82252 epidermal growth f
5	9	2.1	527	2	A42032 epidermal growth f
6	9	2.1	643	2	A36325 epidermal growth f
7	9	2.1	1210	1	GOUHE epidermal growth f
8	9	2.1	1210	1	A53183 epidermal growth f
9	9	2.1	1223	1	TYCHLV carboxyl reductase
10	8	1.9	277	2	JC5284 conserved hypothetical protein
11	8	1.9	341	2	B83298 hypothetical protein
12	8	1.9	463	2	H70922 neuronal cell adhesi
13	8	1.9	725	1	IJMSNG neuronal cell adhesi
14	8	1.9	858	1	IJRTNC neuronal cell adhesi
15	8	1.9	1115	1	IJMSNL hypothetical protein
16	7	1.7	62	2	P90177 hypothetical protein
17	7	1.7	111	2	C35826 hypothetical protein
18	7	1.7	121	2	AH1848 hypothetical protein
19	7	1.7	171	2	F97564 (Y05560) mogg prot
20	7	1.7	171	2	AE2785 hypothetical prote
21	7	1.7	173	2	S10199 NADH dehydrogenas
22	7	1.7	209	2	T0733 hypothetical prote
23	7	1.7	220	2	S05595 trypsin inhibitor
24	7	1.7	220	2	20588 fibroblast growth
25	7	1.7	221	2	JCT7587 stromal cell deriv
26	7	1.7	223	2	H83462 heme exporter prot
27	7	1.7	224	1	A33861 trans-activating t
28	7	1.7	248	2	C83431 type III exptl pr
29	7	1.7	275	2	D75275 endonuclease III -

30	7	1.7	268	2	C70620 hypothesical prote
31	7	1.7	268	2	C90709 thiol disulfide in
32	7	1.7	268	2	G85559 thiol, disulfide in
33	7	1.7	268	2	B64794 hypothetical prote
34	7	1.7	271	2	A45606 DNA-binding prote
35	7	1.7	287	2	G69166 conserved hypothet
36	7	1.7	290	2	T03552 maltose transport
37	7	1.7	301	2	G70613 probable mmaA pro
38	7	1.7	305	2	S47233 catechol 1,2-dioxy
39	7	1.7	306	2	S55233 srp1 protein - Syn
40	7	1.7	307	2	S77651 major membrane pro
41	7	1.7	307	2	T44910 env polyprotein, r
42	7	1.7	312	2	B29350 conserved hypothe
43	7	1.7	312	2	C87552 cation transporter
44	7	1.7	314	2	F86805 hypothetical prote
45	7	1.7	324	2	F97729 hypothetical prote

## ALIGNMENTS

RESULT 1	A24571 protein-tyrosine kinase (EC 2.7.1.112) erbB2 precursor - human
N;Alternate names:	c-erb-B-2 protein precursor; kinase-related transforming protein e
C;Species:	Homo sapiens (man)
C;Date:	Oct-1987 #sequence_revision 06-Dec-1996 #text_change 11-Jun-1999
C;Accession:	A24571; A25491; A44188; B44188; B44188; I59509; 157622
R;Yamamoto, T.; Ikawa, S.; Akiyama, T.; Saito, T.	Nature 319, 230-234, 1986
A;Title:	Similarity of protein encoded by the human c-erb-B-2 gene to epidermal growth
A;Reference number:	A24571; NUID:86118663; PMID:3003577
A;Accession:	A24571
A;Molecule type:	mRNA
A;Residues:	1-1255 <YAM>
A;Cross-references:	GB:03363; NID:93197; PIDN:CAA27060.1; PID:931198
— R;Sembra, K.; Kamata, N.; Toyoshima, K.; Yamamoto, T.	Science 230, 1132-1139, 1985
— proc. Natl. Acad. Sci. U.S.A. 82, 6497-6501, 1985	A;Title: Tyrosine kinase receptor with extensive homology to EGF receptor shares chro
A;Accession:	A25491; NUID:86016729; PMID:2995967
A;Molecule type:	DNA
A;Residues:	737-1031 <SEN>
A;Cross-references:	GB:MI1767; NID:9182163; PIDN:AAA35808.1; PID:9553282
— R;Cousens, L.; Yang-Feng, T.L.; Liao, Y.C.; Chen, E.; McGrath, J.; Seeburg	A;Title: Tyrosine kinase receptor with extensive homology to EGF receptor shares chro
— Science 230, 1132-1139, 1985	A;Reference number: A44188; NUID:8607081; PMID:299974
A;Accession:	A44188
A;Molecule type:	DNA
A;Residues:	740-910 <COOL>
A;Cross-references:	GB:MI2036; NID:9183988; PIDN:AAA35978.1; PID:9183989
A;Accession:	B44188
A;Molecule type:	mRNA
A;Residues:	1-517, RALL, 522, 'S', 524-654, 'V', 656-1169, 'A', 1171-1255 <COU2>
A;Cross-references:	GB:MI1730; NID:913986
— R;King, C.R.; Kraus, M.H.; Aaronson, S.A.	Science 229, 974-976, 1985
A;Title:	Amplification of a novel v-erbB-related gene in a human mammary carcinoma.
A;Reference number:	NUID:85277597; PMID:2992089
A;Accession:	159509
A;Status:	translated from GB/EMBL/DDJB
A;Molecule type:	DNA
A;Residues:	832-909 <REX>
A;Cross-references:	GB:129395; NID:9459807; PIDN:AAA35809.1; PID:9459808
— Mol. Cell. Biol. 7, 2597-2601, 1987	R;Tel, M.; King, C.R.; Kraus, M.H.; Ulrich, A.; Schlessinger, J.; Givol, D.
A;Title:	Human HER2 (neu) promoter: evidence for multiple mechanisms for transcription
A;Reference number:	I57622; NUID:87288898; PMID:3039351
A;Accession:	157622
A;Status:	translated from GB/EMBL/DDJB
A;Molecule type:	DNA
A;Residues:	1-191 <TAL>

A;Cross-references: GB:MI6792; NID:9183993; PIDN:AA58637.1; PID:9553332  
 C;Comment: Amplification and overexpression of this erbB-related gene occurs in about 30% of breast carcinomas.  
 C;Genetics:  
 A;Gene: GDI; ERBB2; NGF; NEU; HBR-2  
 A;Cross-references: GDB:120613; OMIM:164870  
 A;Map position: 17q21.1-17q21.1  
 A;Introns: 25/1; 75/3; 147/1; 883/3  
 A;Note: the list of introns is incomplete  
 C;Function:  
 A;Description: catalyzes the phosphorylation of a peptidyl tyrosine residue by ATP  
 C;Keywords: ATP; autophosphorylation; duplication; glycoprotein; phosphoprotein; phosphoinositide; signal sequence #status predicted <SIG>  
 F;1-21/Domain: protein-tyrosine kinase erbB2 #status predicted <MAT>  
 F;2-1255/Domain: extracellular #status predicted <EXT>  
 F;2-653/Domain: EGF receptor extracellular domain repeat <EX1>  
 F;7-304/Domain: EGF receptor extracellular domain repeat <EX2>  
 F;395-605/Domain: transmembrane #status predicted <TM>  
 F;654-675/Domain: intracellular #status predicted <INT>  
 F;7-18-983/Domain: protein kinase homology <KIN>  
 F;7-26-734/Region: protein kinase ATP-binding motif  
 F;68-124-187, 259 530-571-629/Binding site: carbohydrate (Asn) (covalent) #status predicted <CAR>  
 F;686/RBinding site: phosphate (Thr) (covalent) (by Protein kinase C) #status predicted <PROT>  
 F;1139,1221-1222,1248/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation)  
 Query Match 19.8%; Score 83; DB 1; Length 1255;  
 Best Local Similarity 100.0%; Pred. No. 2.9e-74; Mismatches 0; Indels 0; Gaps 0;  
 Matches 83; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 199 CKGSRQWGESSEDCOSLRTVCGGSCARCKPLPTPCCHEQCAAGCTGPKHSDCLACLFH 258  
 Db 199 CKGSRQWGESSEDCOSLRTVCGGCARCKPLPTDCCHEQCAAGCTGPKHSDCLACLFH 258  
 Qy 259 NHSGICELHCALVITYNTDFES 281  
 Db 259 NHSGICELHCALVITYNTDFES 281

RESULT 2  
 148161  
 p-105 precursor - golden hamster (golden hamster)  
 C;Species: Mesocricetus auratus (golden hamster)  
 C;Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 18-Jun-1999  
 C;Accession: I48161  
 R;Nakanura, T.; Ushijima, T.; Ishizaka, Y.; Nagao, M.; Arai, M.; Yamazaki, Y.; Ishikawa,  
 A;Title: Cloning and activation of the Syrian hamster neu proto-oncogene.  
 A;Reference number: I48161; MUID:94193007; PMID:7908275  
 A;Accession: I48161  
 A;Status: Preliminary; translated from GB/EMBL/DDJB  
 A;Molecule type: mRNA  
 A;Residues: 1-1254 <RES>  
 A;Cross-references: GB:DI6295; NID:9493236; PIDN:BA03801.1; PID:9747595  
 A;Gene: neu  
 C;Superfamily: epidermal growth factor receptor; protein kinase homology  
 C;Keywords: ATP  
 F;7-18-983/Domain: protein kinase homology <KIN>  
 F;7-26-734/Region: protein kinase ATP-binding motif

Query Match 12.2%; Score 51; DB 1; Length 1260;  
 Best Local Similarity 100.0%; Pred. No. 3e-42; Mismatches 0; Indels 0; Gaps 0;  
 Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 231 LPTDCCHEQCAAGCTGPKHSDCLACLFHNHSIGICELHCALVITYNTDFES 281  
 Db 235 LPTDCCHEQCAAGCTGPKHSDCLACLFHNHSIGICELHCALVITYNTDFES 281

RESULT 4  
 882252  
 Rnf-related protein Vc1012 [Imported] - Vibrio cholerae (strain N16961 serogroup O1)  
 C;Species: Vibrio cholerae  
 C;Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 02-Feb-2001  
 C;Accession: B82252  
 R;Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.; Chardson, D.; Embley, M.D.; Vamathevan, J.; Basa, S.; Qin, H.; Dragoi, I.; Sellers, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.  
 Nature 406, 477-483, 2000  
 A;Title: DNA sequence of both chromosomes of the cholera pathogen Vibrio cholerae.  
 A;Reference number: A82035; MUID:20406833; PMID:10952301  
 A;Accession: B82252  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-230 <RES>  
 A;Cross-references: GB:AE004183; GB:AE003852; NID:99655473; PIDN:AAF94173.1; GSFD:GN  
 A;Experimental source: serogroup O1; strain N16961; biotype El Tor  
 C;Genetics:  
 A;Gene: VC1012  
 A;Map Position: 1  
 C;Superfamily: conserved hypothetical protein H1688

Query Match 2.1%; Score 9; DB 2; Length 230;  
 Best Local Similarity 100.0%; Pred. No. 0.7; Mismatches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 231 LPTDCCHEQCAAGCTGPKHSDCLACLFHNHSIGICELHCALVITYNTDFES 281  
 Db 231 LPTDCCHEQCAAGCTGPKHSDCLACLFHNHSIGICELHCALVITYNTDFES 281

RESULT 3  
 TVRING

Qy 12 LLALLPPGA 20

Db 184 LLAUPLPPGA 192  
 A;Molecule type: mRNA  
 A;Residues: 1-1210 <ULL>  
 A;Cross-references: EMBL:X00588; NID:931113; PIDN:CAA25240.1; PID:g757924  
 C;Species: Gallus gallus (chicken)  
 C;Date: 31-Dec-1993 #sequence\_revision 31-dec-1993 #text\_change 18-Jun-1999  
 C;Accession: A42032  
 C;Accession: A42032  
 R;Flickinger, T.W.; Maibl, N.J.; Kung, H.J.  
 Mol. Cell. Biol. 12, 883-893, 1992  
 A;Title: An alternative processed mRNA from the avian c-erbB gene encodes a soluble, t  
 A;Cross-references: GB:W77637; NID:9211737; PIDN:AAA8759.1; PID:9211738  
 A;Experimental source: liver  
 A;Note: sequence extracted from NCBI backbone (NCBIN:76892, NCBIPI:76893)  
 C;Superfamily: epidermal growth factor receptor; protein kinase homology  
 C;Keywords: ATP; growth factor receptor  
 A;Residues: 1-527 <FLI>  
 Query Match 2.1%; Score 9; DB 2; Length 527;  
 Best Local Similarity 100.0%; Pred. No. 1.4;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Db 239 QCAAGCTGP 247  
 |||||||  
 Db 245 QCAAGCTGP 253  
 |||||||  
**RESULT 6**  
 A36325  
 epidermal growth factor receptor - rat  
 C;Species: Rattus norvegicus (Norway rat)  
 C;Date: 25-Jan-1991 #sequence\_revision 25-Jan-1991 #text\_change 10-Oct-1997  
 C;Accession: A36325  
 R;Petch, L.A.; Harris, J.; Raymond, V.W.; Blasband, A.; Lee, D.C.; Earp, H.S.  
 Mol. Cell. Biol. 10, 2973-2982, 1990  
 A;Title: A truncated, secreted form of the epidermal growth factor receptor is encoded by  
 A;Reference number: A36325; MUID:90258888; PMID:2342466  
 A;Status: preliminary  
 A;Molecule type: mRNA  
 A;Residues: 1-644 <PET>  
 A;Cross references: GB:M37394  
 C;Superfamily: epidermal growth factor receptor; protein kinase homology  
 C;Keywords: alternative splicing; ATP; growth factor receptor  
 Query Match 2.1%; Score 9; DB 2; Length 644;  
 Best Local Similarity 100.0%; Pred. No. 1.7;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Db 239 QCAAGCTGP 247  
 |||||||  
 Db 235 QCAAGCTGP 243  
 |||||||  
**RESULT 7**  
 GQ10E  
 epidermal growth factor receptor precursor - human  
 N;Contains: protein-tyrosine kinase (EC 2.7.1.112) erbB  
 C;Species: Homo sapiens (man)  
 C;Date: 15-Nov-1984 #sequence\_revision 27-Nov-1985 #text\_change 11-Jun-1999  
 C;Accession: A00641; A25772; S30024; A38672; A00642; A43615; A23062; A05081; A60143; A33  
 C;Accession: A00641; A25772; S30024; A38672; A00642; A43615; A23062; A05081; A60143; A33  
 R;Ullrich, A.; Coussens, L.; Hayflick, J.S.; Dull, T.J.; Gray, A.; Tam, A.W.; Lee, J.; Y  
 Nature 309, 418-425, 1984  
 A;Title: Human epidermal growth factor receptor cDNA sequence and aberrant expression of  
 A;Reference number: A00641; MUID:84219729; PMID:632812  
 A;Accession: A00641  
 A;Molecule type: DNA  
 A;Residues: 1-29 <ISH>  
 A;Cross-references: GB:MI1234; NID:918981; PIDN:AAA2370.1; PID:9533272  
 A;Residues: 1-29 <RA2>  
 A;Cross-references: EMBL:X06370; NID:931118; PIDN:CAA29668.1; PID:931119  
 R;Haley, J.J.; Whittle, N.; Bennett, P.; Kinchington, D.; Ulrich, A.; Waterfield, M.  
 Oncogene, Res., 1, 315-316, 1987  
 A;Title: The human EGF receptor gene: structure of the 110 kb locus and identification  
 A;Reference number: S30024; MUID:88217333; PMID:3329716  
 A;Accession: S30024  
 A;Molecule type: DNA  
 A;Residues: 1-29 <HAL>  
 A;Cross-references: GB:M38425; NID:9181977; PIDN:AAA63171.1; PID:9553271  
 A;Experimental source: carcinoma cell line A431/  
 A;Molecule type: DNA  
 A;Residues: 1-29 <HAL>  
 A;Cross-references: GB:M38425; NID:9181977; PIDN:AAA63171.1; PID:9553271  
 A;Experimental source: carcinoma cell line A431/  
 R;Xu, Y.; Ishii, S.; Clark, A.J.L.; Sullivan, M.; Wilson, R.K.; Ma, D.P.; Roe, B.A.;  
 Nature 309, 806-810, 1984  
 A;Title: Human epidermal growth factor receptor cDNA is homologous to a variety of RN  
 A;Reference number: A00642; MUID:84245835; PMID:6330563  
 A;Accession: A00642  
 A;Molecule type: mRNA  
 A;Residues: RCWAWR, 150-187, 'KSYTOAV', 195, 'M' 197, 'A', 199-222, 'S', 224-304, 'R', 307-  
 'T', 308-311, 'D', 312-315, 'P', 316-319, 'R', 320-323, 'Y', 324-327, 'V', 328-331, 'A', 332-335, 'G', 336-339, 'C', 340-343, 'F', 344-347, 'W', 348-351, 'L', 352-355, 'I', 356-359, 'V', 360-363, 'T', 364-367, 'S', 368-371, 'N', 372-375, 'P', 376-379, 'A', 380-383, 'G', 384-387, 'C', 388-391, 'D', 392-395, 'E', 396-399, 'R', 400-403, 'K', 404-407, 'Q', 408-411, 'H', 412-415, 'S', 416-419, 'N', 420-423, 'T', 424-427, 'Y', 428-431, 'F', 432-435, 'W', 436-439, 'L', 440-443, 'I', 444-447, 'V', 448-451, 'A', 452-455, 'G', 456-459, 'C', 460-463, 'D', 464-467, 'E', 468-471, 'R', 472-475, 'K', 476-479, 'Q', 480-483, 'H', 484-487, 'S', 488-491, 'N', 492-495, 'P', 496-499, 'A', 500-503, 'G', 504-507, 'C', 508-511, 'D', 512-515, 'E', 516-519, 'R', 520-523, 'K', 524-527, 'Q', 528-531, 'H', 532-535, 'S', 536-539, 'N', 540-543, 'P', 544-547, 'A', 548-551, 'G', 552-555, 'C', 556-559, 'D', 560-563, 'E', 564-567, 'R', 568-571, 'K', 572-575, 'Q', 576-579, 'H', 580-583, 'S', 584-587, 'N', 588-591, 'P', 592-595, 'A', 596-599, 'G', 600-603, 'C', 604-607, 'D', 608-611, 'E', 612-615, 'R', 616-619, 'K', 620-623, 'Q', 624-627, 'H', 628-631, 'S', 632-635, 'N', 636-639, 'P', 640-643, 'A', 644-647, 'G', 648-651, 'C', 652-655, 'D', 656-659, 'E', 660-663, 'R', 664-667, 'K', 668-671, 'Q', 672-675, 'H', 676-679, 'S', 680-683, 'N', 684-687, 'P', 688-691, 'A', 692-695, 'G', 696-699, 'C', 700-703, 'D', 704-707, 'E', 708-711, 'R', 712-715, 'K', 716-719, 'Q', 720-723, 'H', 724-727, 'S', 728-731, 'N', 732-735, 'P', 736-739, 'A', 740-743, 'G', 744-747, 'C', 748-751, 'D', 752-755, 'E', 756-759, 'R', 760-763, 'K', 764-767, 'Q', 768-771, 'H', 772-775, 'S', 776-779, 'N', 780-783, 'P', 784-787, 'A', 788-791, 'G', 792-795, 'C', 796-799, 'D', 800-803, 'E', 804-807, 'R', 808-811, 'K', 812-815, 'Q', 816-819, 'H', 820-823, 'S', 824-827, 'N', 828-831, 'P', 832-835, 'A', 836-839, 'G', 840-843, 'C', 844-847, 'D', 848-851, 'E', 852-855, 'R', 856-859, 'K', 860-863, 'Q', 864-867, 'H', 868-871, 'S', 872-875, 'N', 876-879, 'P', 880-883, 'A', 884-887, 'G', 888-891, 'C', 892-895, 'D', 896-899, 'E', 900-903, 'R', 904-907, 'K', 908-911, 'Q', 912-915, 'H', 916-919, 'S', 920-923, 'N', 924-927, 'P', 928-931, 'A', 932-935, 'G', 936-939, 'C', 940-943, 'D', 944-947, 'E', 948-951, 'R', 952-955, 'K', 956-959, 'Q', 960-963, 'H', 964-967, 'S', 968-971, 'N', 972-975, 'P', 976-979, 'A', 980-983, 'G', 984-987, 'C', 988-991, 'D', 992-995, 'E', 996-999, 'R', 1000-1003, 'K', 1004-1007, 'Q', 1008-1011, 'H', 1012-1015, 'S', 1016-1019, 'N', 1020-1023, 'P', 1024-1027, 'A', 1028-1031, 'G', 1032-1035, 'C', 1036-1039, 'D', 1040-1043, 'E', 1044-1047, 'R', 1048-1051, 'K', 1052-1055, 'Q', 1056-1059, 'H', 1060-1063, 'S', 1064-1067, 'N', 1068-1071, 'P', 1072-1075, 'A', 1076-1079, 'G', 1080-1083, 'C', 1084-1087, 'D', 1088-1091, 'E', 1092-1095, 'R', 1096-1099, 'K', 1100-1103, 'Q', 1104-1107, 'H', 1108-1111, 'S', 1112-1115, 'N', 1116-1119, 'P', 1120-1123, 'A', 1124-1127, 'G', 1128-1131, 'C', 1132-1135, 'D', 1136-1139, 'E', 1140-1143, 'R', 1144-1147, 'K', 1148-1151, 'Q', 1152-1155, 'H', 1156-1159, 'S', 1160-1163, 'N', 1164-1167, 'P', 1168-1171, 'A', 1172-1175, 'G', 1176-1179, 'C', 1180-1183, 'D', 1184-1187, 'E', 1188-1191, 'R', 1192-1195, 'K', 1196-1199, 'Q', 1200-1203, 'H', 1204-1207, 'S', 1208-1211, 'N', 1212-1215, 'P', 1216-1219, 'A', 1220-1223, 'G', 1224-1227, 'C', 1228-1231, 'D', 1232-1235, 'E', 1236-1239, 'R', 1240-1243, 'K', 1244-1247, 'Q', 1248-1251, 'H', 1252-1255, 'S', 1256-1259, 'N', 1260-1263, 'P', 1264-1267, 'A', 1268-1271, 'G', 1272-1275, 'C', 1276-1279, 'D', 1280-1283, 'E', 1284-1287, 'R', 1288-1291, 'K', 1292-1295, 'Q', 1296-1299, 'H', 1300-1303, 'S', 1304-1307, 'N', 1308-1311, 'P', 1312-1315, 'A', 1316-1319, 'G', 1320-1323, 'C', 1324-1327, 'D', 1328-1331, 'E', 1332-1335, 'R', 1336-1339, 'K', 1340-1343, 'Q', 1344-1347, 'H', 1348-1351, 'S', 1352-1355, 'N', 1356-1359, 'P', 1360-1363, 'A', 1364-1367, 'G', 1368-1371, 'C', 1372-1375, 'D', 1376-1379, 'E', 1380-1383, 'R', 1384-1387, 'K', 1388-1391, 'Q', 1392-1395, 'H', 1396-1399, 'S', 1400-1403, 'N', 1404-1407, 'P', 1408-1411, 'A', 1412-1415, 'G', 1416-1419, 'C', 1420-1423, 'D', 1424-1427, 'E', 1428-1431, 'R', 1432-1435, 'K', 1436-1439, 'Q', 1440-1443, 'H', 1444-1447, 'S', 1448-1451, 'N', 1452-1455, 'P', 1456-1459, 'A', 1460-1463, 'G', 1464-1467, 'C', 1468-1471, 'D', 1472-1475, 'E', 1476-1479, 'R', 1480-1483, 'K', 1484-1487, 'Q', 1488-1491, 'H', 1492-1495, 'S', 1496-1499, 'N', 1500-1503, 'P', 1504-1507, 'A', 1508-1511, 'G', 1512-1515, 'C', 1516-1519, 'D', 1520-1523, 'E', 1524-1527, 'R', 1528-1531, 'K', 1532-1535, 'Q', 1536-1539, 'H', 1540-1543, 'S', 1544-1547, 'N', 1548-1551, 'P', 1552-1555, 'A', 1556-1559, 'G', 1560-1563, 'C', 1564-1567, 'D', 1568-1571, 'E', 1572-1575, 'R', 1576-1579, 'K', 1580-1583, 'Q', 1584-1587, 'H', 1588-1591, 'S', 1592-1595, 'N', 1596-1599, 'P', 1600-1603, 'A', 1604-1607, 'G', 1608-1611, 'C', 1612-1615, 'D', 1616-1619, 'E', 1620-1623, 'R', 1624-1627, 'K', 1628-1631, 'Q', 1632-1635, 'H', 1636-1639, 'S', 1640-1643, 'N', 1644-1647, 'P', 1648-1651, 'A', 1652-1655, 'G', 1656-1659, 'C', 1660-1663, 'D', 1664-1667, 'E', 1668-1671, 'R', 1672-1675, 'K', 1676-1679, 'Q', 1680-1683, 'H', 1684-1687, 'S', 1688-1691, 'N', 1692-1695, 'P', 1696-1699, 'A', 1700-1703, 'G', 1704-1707, 'C', 1708-1711, 'D', 1712-1715, 'E', 1716-1719, 'R', 1720-1723, 'K', 1724-1727, 'Q', 1728-1731, 'H', 1732-1735, 'S', 1736-1739, 'N', 1740-1743, 'P', 1744-1747, 'A', 1748-1751, 'G', 1752-1755, 'C', 1756-1759, 'D', 1760-1763, 'E', 1764-1767, 'R', 1768-1771, 'K', 1772-1775, 'Q', 1776-1779, 'H', 1780-1783, 'S', 1784-1787, 'N', 1788-1791, 'P', 1792-1795, 'A', 1796-1799, 'G', 1800-1803, 'C', 1804-1807, 'D', 1808-1811, 'E', 1812-1815, 'R', 1816-1819, 'K', 1820-1823, 'Q', 1824-1827, 'H', 1828-1831, 'S', 1832-1835, 'N', 1836-1839, 'P', 1840-1843, 'A', 1844-1847, 'G', 1848-1851, 'C', 1852-1855, 'D', 1856-1859, 'E', 1860-1863, 'R', 1864-1867, 'K', 1868-1871, 'Q', 1872-1875, 'H', 1876-1879, 'S', 1880-1883, 'N', 1884-1887, 'P', 1888-1891, 'A', 1892-1895, 'G', 1896-1899, 'C', 1900-1903, 'D', 1904-1907, 'E', 1908-1911, 'R', 1912-1915, 'K', 1916-1919, 'Q', 1920-1923, 'H', 1924-1927, 'S', 1928-1931, 'N', 1932-1935, 'P', 1936-1939, 'A', 1940-1943, 'G', 1944-1947, 'C', 1948-1951, 'D', 1952-1955, 'E', 1956-1959, 'R', 1960-1963, 'K', 1964-1967, 'Q', 1968-1971, 'H', 1972-1975, 'S', 1976-1979, 'N', 1980-1983, 'P', 1984-1987, 'A', 1988-1991, 'G', 1992-1995, 'C', 1996-1999, 'D', 2000-2003, 'E', 2004-2007, 'R', 2008-2011, 'K', 2012-2015, 'Q', 2016-2019, 'H', 2020-2023, 'S', 2024-2027, 'N', 2028-2031, 'P', 2032-2035, 'A', 2036-2039, 'G', 2040-2043, 'C', 2044-2047, 'D', 2048-2051, 'E', 2052-2055, 'R', 2056-2059, 'K', 2060-2063, 'Q', 2064-2067, 'H', 2068-2071, 'S', 2072-2075, 'N', 2076-2079, 'P', 2080-2083, 'A', 2084-2087, 'G', 2088-2091, 'C', 2092-2095, 'D', 2096-2099, 'E', 2100-2103, 'R', 2104-2107, 'K', 2108-2111, 'Q', 2112-2115, 'H', 2116-2119, 'S', 2120-2123, 'N', 2124-2127, 'P', 2128-2131, 'A', 2132-2135, 'G', 2136-2139, 'C', 2140-2143, 'D', 2144-2147, 'E', 2148-2151, 'R', 2152-2155, 'K', 2156-2159, 'Q', 2160-2163, 'H', 2164-2167, 'S', 2168-2171, 'N', 2172-2175, 'P', 2176-2179, 'A', 2180-2183, 'G', 2184-2187, 'C', 2188-2191, 'D', 2192-2195, 'E', 2196-2199, 'R', 2200-2203, 'K', 2204-2207, 'Q', 2208-2211, 'H', 2212-2215, 'S', 2216-2219, 'N', 2220-2223, 'P', 2224-2227, 'A', 2228-2231, 'G', 2232-2235, 'C', 2236-2239, 'D', 2240-2243, 'E', 2244-2247, 'R', 2248-2251, 'K', 2252-2255, 'Q', 2256-2259, 'H', 2260-2263, 'S', 2264-2267, 'N', 2268-2271, 'P', 2272-2275, 'A', 2276-2279, 'G', 2280-2283, 'C', 2284-2287, 'D', 2288-2291, 'E', 2292-2295, 'R', 2296-2299, 'K', 2300-2303, 'Q', 2304-2307, 'H', 2308-2311, 'S', 2312-2315, 'N', 2316-2319, 'P', 2320-2323, 'A', 2324-2327, 'G', 2328-2331, 'C', 2332-2335, 'D', 2336-2339, 'E', 2340-2343, 'R', 2344-2347, 'K', 2348-2351, 'Q', 2352-2355, 'H', 2356-2359, 'S', 2360-2363, 'N', 2364-2367, 'P', 2368-2371, 'A', 2372-2375, 'G', 2376-2379, 'C', 2380-2383, 'D', 2384-2387, 'E', 2388-2391, 'R', 2392-2395, 'K', 2396-2399, 'Q', 2400-2403, 'H', 2404-2407, 'S', 2408-2411, 'N', 2412-2415, 'P', 2416-2419, 'A', 2420-2423, 'G', 2424-2427, 'C', 2428-2431, 'D', 2432-2435, 'E', 2436-2439, 'R', 2440-2443, 'K', 2444-2447, 'Q', 2448-2451, 'H', 2452-2455, 'S', 2456-2459, 'N', 2460-2463, 'P', 2464-2467, 'A', 2468-2471, 'G', 2472-2475, 'C', 2476-2479, 'D', 2480-2483, 'E', 2484-2487, 'R', 2488-2491, 'K', 2492-2495, 'Q', 2496-2499, 'H', 2500-2503, 'S', 2504-2507, 'N', 2508-2511, 'P', 2512-2515, 'A', 2516-2519, 'G', 2520-2523, 'C', 2524-2527, 'D', 2528-2531, 'E', 2532-2535, 'R', 2536-2539, 'K', 2540-2543, 'Q', 2544-2547, 'H', 2548-2551, 'S', 2552-2555, 'N', 2556-2559, 'P', 2560-2563, 'A', 2564-2567, 'G', 2568-2571, 'C', 2572-2575, 'D', 2576-2579, 'E', 2580-2583, 'R', 2584-2587, 'K', 2588-2591, 'Q', 2592-2595, 'H', 2596-2599, 'S', 2600-2603, 'N', 2604-2607, 'P', 2608-2611, 'A', 2612-2615, 'G', 2616-2619, 'C', 2620-2623, 'D', 2624-2627, 'E', 2628-2631, 'R', 2632-2635, 'K', 2636-2639, 'Q', 2640-2643, 'H', 2644-2647, 'S', 2648-2651, 'N', 2652-2655, 'P', 2656-2659, 'A', 2660-2663, 'G', 2664-2667, 'C', 2668-2671, 'D', 2672-2675, 'E', 2676-2679, 'R', 2680-2683, 'K', 2684-2687, 'Q', 2688-2691, 'H', 2692-2695, 'S', 2696-2699, 'N', 2700-2703, 'P', 2704-2707, 'A', 2708-2711, 'G', 2712-2715, 'C', 2716-2719, 'D', 2720-2723, 'E', 2724-2727, 'R', 2728-2731, 'K', 2732-2735, 'Q', 2736-2739, 'H', 2740-2743, 'S', 2744-2747, 'N', 2748-2751, 'P', 2752-2755, 'A', 2756-2759, 'G', 2760-2763, 'C', 2764-2767, 'D', 2768-2771, 'E', 2772-2775, 'R', 2776-2779, 'K', 2780-2783, 'Q', 2784-2787, 'H', 2788-2791, 'S', 2792-2795, 'N', 2796-2799, 'P', 2800-2803, 'A', 2804-2807, 'G', 2808-2811, 'C', 2812-2815, 'D', 2816-2819, 'E', 2820-2823, 'R', 2824-2827, 'K', 2828-2831, 'Q', 2832-2835, 'H', 2836-2839, 'S', 2840-2843, 'N', 2844-2847, 'P', 2848-2851, 'A', 2852-2855, 'G', 2856-2859, 'C', 2860-2863, 'D', 2864-2867, 'E', 2868-2871, 'R', 2872-2875, 'K', 2876-2879, 'Q', 2880-2883, 'H', 2884-2887, 'S', 2888-2891, 'N', 2892-2895, 'P', 2896-2899, 'A', 2900-2903, 'G', 2904-2907, 'C', 2908-2911, 'D', 2912-2915, 'E', 2916-2919, 'R', 2920-2923, 'K', 2924-2927, 'Q', 2928-2931, 'H', 2932-2935, 'S', 2936-2939, 'N', 2940-2943, 'P', 2944-2947, 'A', 2948-2951, 'G', 2952-2955, 'C', 2956-2959, 'D', 2960-2963, 'E', 2964-2967, 'R', 2968-2971, 'K', 2972-2975, 'Q', 2976-2979, 'H', 2980-2983, 'S', 2984-2987, 'N', 2988-2991, 'P', 2992-2995, 'A', 2996-2999, 'G', 3000-3003, 'C', 3004-3007, 'D', 3008-3011, 'E', 3012-3015, 'R', 3016-3019, 'K', 3020-3023, 'Q', 3024-3027

Cell 59, 33-43, 1989  
A;Title: Functional independence of the epidermal growth factor receptor from a domain  
A;Reference number: A33331; MUID:9003233; PMID:2790960  
A;Map Position: 7p12.3-7p12.1  
C;Superfamily: epidermal growth factor receptor; internalization signal  
C;Keywords: ATP; autophosphorylation; duplication; glycoprotein; phosphoprotein; phospho  
F;1-24/Domain: signal sequence #status predicted <SIG>  
F;25-1210/Product: EGF receptor #status predicted <EXT>  
F;75-300/Domain: protein kinase homology <KIN>  
F;999-106/Region: coated pit mediated internalization signal  
F;104-1210/Region: inhibitory  
F;66-668/binding: transmembrane #status predicted <TMM>  
F;669-1210/Domain: intracellular #status predicted <INT>  
F;710-975/Domain: protein kinase homology <KIN>  
F;99-726/Region: protein kinase ATP-binding motif  
F;104-1210/Region: inhibitory  
F;175-352,413,444-528/603/Binding site: carbohydrate (Asn) (covalent) #status predicted <TMM>  
F;745/Active site: Lys #status experimental  
Query Match 2.1%; Score 9; DB 1; Length 1210;  
Best Local Similarity 100.0%; Pred. No. 3;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 239 QCAAGCTGP 247  
Db 235 QCAAGCTGP 243

**RESULT 8**

A53183  
epidermal growth factor receptor precursor - mouse  
C;Species: Mus musculus (house mouse)  
C;Date: 06-Jan-1995 #sequence\_revision 05-Jan-1995 #text\_change 18-Jun-1999  
C;Accession: A53183; A43818; S24942; A2841; S45325; I45643  
R;Luettekemper, N.C.; Phillips, H.K.; Qu, T.H.; Copeland, N.G.; Earp, H.S.; Jenkins, N.A.;  
Genes Dev., 8, 399-413, 1994  
A;Title: The mouse waved-2 phenotype results from a point mutation in the EGF receptor t  
A;Reference number: A53183; MUID:94170986; PMID:8125255  
A;Accession: A53183  
A;Molecule type: mRNA  
A;Residues: 1-1210 <JUN>  
A;Cross-references: GB:U03425  
R;Aviñó, A.; Lax, I.; Ulrich, A.; Schlessinger, J.; Givol, D.; Morse, B.  
Oncogene 6, 673-676, 1991  
A;Title: Comparison of EGF receptor sequences as a guide to study the ligand binding sit  
A;Reference number: A43818; MUID:91232866; PMID:2030916  
A;Accession: A43818  
A;Molecule type: mRNA  
A;Residues: 1-714 <AVI>  
A;Cross-references: GB:X59698  
R;Eisinger, D.P.; Segreto, G.  
submitted to the EMBL Data Library, June 1992  
A;Reference number: S24942  
A;Accession: S24942  
A;Molecule type: mRNA  
A;Residues: 969-971, 'K', 973-1115, 'D' <EIS>  
A;Cross-references: EMBL:212608  
R;Heissmann, G.J.; Gill, G.N.  
J. Biol. Chem., 263, 13152-13158, 1988  
A;Title: Epidermal growth factor receptor threonine and serine residues phosphorylated ;  
A;Reference number: A28941; MUID:88330814; PMID:3138233  
A;Accession: A28941  
A;Molecule type: protein  
R;Residues: 689-694, 'X', 696-704, 'L', 706-707; 989-992, 'XX', 995-996, 'X', 998-1000; 1002-1009,  
submitted to the EMBL Data Library, April 1994  
A;Description: The complete cDNA sequence of the Mouse Epidermal Growth Factor Receptor

A;Reference number: S45325  
A;Accession: S45325  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-971, 'K', 973-1210 <VER>  
A;Cross-references: EMBL:X8987; NID:9488830; PIDN:CAA55587\_1; PID:9488831  
R;Paria, B.C.; Das, S.K.; Andrews, G.K.; Dey, S.K.  
Proc. Natl. Acad. Sci. U.S.A. 90, 55-59, 1993  
A;Title: Expression of the epidermal growth factor receptor gene is regulated in mous  
A;Reference number: 149643; MUID:93126380; PMID:7678348  
A;Accession: 149643  
A;Status: translated from GB/EMBL/DDBJ  
A;Molecule type: mRNA  
A;Residues: 12-20, 132 <RES>  
A;Cross-references: GB:L06864; NID:9193001; PIDN:AA53029\_1; PID:9567201  
C;Genetics:  
A;Gene: EGFR  
C;Superfamily: epidermal growth factor receptor; protein kinase homology  
C;Keywords: ATP; growth factor receptor; kinase-related transforming protein; phospho  
F;1-24/Domain: signal sequence #status predicted <SIG>  
F;1648-670/Domain: transmembrane #status predicted <TMM>  
F;712-977/Domain: protein kinase homology <KIN>  
F;720-728/Region: protein kinase ATP-binding motif  
F;680-695/Binding site: phosphate (Thr) (covalent) #status experimental  
F;697-1070,1071/Binding site: Phosphate (Ser) (covalent) #status experimental  
F;1028/Binding site: (or 997) phosphate (Ser) (covalent) #status experimental  
F;1197/Binding site: phosphate (Tyr) (covalent) #status experimental  
Query Match 2.1%; Score 9; DB 2; Length 1210;  
Best Local Similarity 100.0%; Pred. No. 3;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 239 QCAAGCTGP 247  
Db 235 QCAAGCTGP 243

**RESULT 9**

TVCHLV  
epidermal growth factor receptor precursor - chicken  
C;Contains: protein-tyrosine kinase (EC 2.7.1.112) erbB  
C;Species: Gallus gallus (chicken)  
C;Date: 28-Feb-1986 #sequence\_revision 05-May-1995 #text\_change 04-Feb-2000  
C;Accession: A27720; R00643  
R;Lax, I.; Johnson, A.; Howk, R.; Sap, J.; Bellot, F.; Winkler, M.; Ulrich, A.; Venn  
Mol. Cell. Biol., 8, 1970-1978, 1988  
A;Title: Chicken epidermal growth factor (EGF) receptor: cDNA cloning, expression in  
A;Reference number: A27720; MUID:88261272; PMID:3260329  
A;Accession: A27720  
A;Molecule type: mRNA  
A;Residues: 1-1223 <LAX>  
A;Cross-references: GB:M20386  
R;Nilsen, T.W.; Maroney, P.A.; Goodwin, R.G.; Rottman, F.M.; Crittenden, L.B.; Raines  
Cell 41, 719-726, 1985  
A;Title: c-erbB activation in ALV-induced erythroblastosis: novel RNA processing and  
A;Reference number: A00643; MUID:85228222; PMID:2988784  
A;Accession: A00643  
A;Molecule type: mRNA  
A;Residues: 585-1223 <NTL>  
A;Cross-references: GB:M1066  
A;Gene: rbb  
C;Superfamily: epidermal growth factor receptor; protein kinase homology  
C;Keywords: alternative splicing; ATP; autophosphorylation; glycoprotein; growth fact  
specific Protein kinase  
F;1-30/Domain: signal sequence #status predicted <SIG>  
F;31-1223/Product: epidermal growth factor receptor #status predicted <MAT>  
F;31-654/Domain: extracellular #status predicted <EXT>  
F;181-307/Domain: EGF receptor extracellular domain repeat <EE1>  
F;397-610/Domain: EGF receptor extracellular domain repeat <EE2>  
F;655-677/Domain: transmembrane #status predicted <TMM>  
F;678-1223/Domain: intracellular #status predicted <INT>

F;719-984#domain: protein kinase homology <IN>  
 F;727-735#Region: protein kinase ATP-binding motif  
 F;136,202,280,361,370,422,575,580,615,635#Binding site: carbohydrate (Thr) (covalent) #status predicted  
 F;192,650#Binding site: carbohydrate (Ser) (covalent) #status predicted  
 F;687#Active site: phosphate (Tyr) (covalent) (by protein kinase C) #status predicted  
 F;1100,1183,1208#Binding site: phosphate (Tyr) (covalent) (by autophosphorylation) #status predicted  
 ; Lory, S.; Olson, M.V.  
 Query Match 2.1%; Score 9; DB 1; Length 1223;  
 Best Local Similarity 100.0%; Pred. No. 3;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 239 QCAAGCTGP 247  
 Db 242 QCAGAGCTGP 250

**RESULT 10**

JC2284 carbonyl reductase (NADPH2) (EC 1.1.1.184), inducible - rat  
 C;Species: Rattus norvegicus (Norway rat)  
 C;Date: 16-Apr-1997 #sequence\_revision 18-Jul-1997 #text\_change 03-Jun-2002  
 C;Accession: S68982; P2234; JC5284; PC2159; S52349  
 R;Hermuth, B.; Maeder-Hainemann, G.; Ernst, E.  
 Eur. J. Biochem. 228, 473-479, 1995  
 A;Title: Cloning and expression of carbonyl reductase from rat testis.  
 A;Reference number: S68982; MUID:95220378; PMID:7705364

A;Status: preliminary  
 A;Molecule type: mRNA  
 A;Residues: 1-277 <WERS>  
 A;Cross-references: EMBL:X84349; NID:9666086; PIDN:CAA59088.1; PID:9666087  
 R;Toft, E.; Soederstrom, M.; Ahlberg, M.B.; Depierre, J.W.  
 Biochem. Biophys. Res. Commun. 201, 149-154, 1994  
 A;Title: A novel 34kDa glutathione-binding protein in mature rat ovary.  
 A;Reference number: PC2159; MUID:94256971; PMID:8198567

A;Accession: PC2234  
 A;Molecule type: protein  
 A;Residues: 104-121, X, 123-134, D', 136-137, H' <TOF>  
 A;Experimental source: ovary  
 R;Aoki, H.; Okada, T.; Mizutani, T.; Numata, Y.; Minegishi, T.; Miyamoto, K.  
 Biochem. Biophys. Res. Commun. 230, 518-523, 1997  
 A;Title: Identification of two closely related genes, inducible and noninducible carbonyl  
 A;Reference number: JC2284; MUID:97167735; PMID:901553  
 A;Contents: ovary  
 A;Accession: JC5284  
 A;Molecule type: mRNA  
 A;Residues: 1-140, 'GM', 143, 'R', 145-235, 'T', 237-238, 'E', 240-277 <AOK>  
 A;Cross-references: DBJ:D8069; NID:91906311; PIDN:BA1907.1; PID:9190812  
 C;Comment: this enzyme is a cytosolic, monomeric oxidoreductase that catalyzes the NADP+  
 C;Genetics:  
 A;Gene: ICR  
 C;SuperFamily: ribitol dehydrogenase; short-chain alcohol dehydrogenase homology  
 C;Keywords: cytosol; monomer; NADP; oxidoreductase  
 F;6-185#domain: short-chain alcohol dehydrogenase homology <SDH>  
 Query Match 1.9%; Score 8; DB 2; Length 277;  
 Best Local Similarity 100.0%; Pred. No. 8.2;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 13 LALUUPGA 20  
 Db 254 LALUUPGA 261

**RESULT 11**

B83298 conserved hypothetical protein Ph2771 [imported] - Pseudomonas aeruginosa (strain PA01)  
 C;Species: Pseudomonas aeruginosa  
 C;Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 31-Dec-2000  
 C;Accession: B83298  
 R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizooguchi, S.D.; Warren, P.; Hickey, M.J.; Braman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Harbig, K.; Lim, J.

Nature 406, 959-964, 2000  
 A;Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic pa  
 A;Reference number: A82950; MUID:20437337; PMID:10984043  
 A;Accession: B83298  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-341 <STO>  
 A;Cross-references: GB:AB004705; GB:AE004091; NID:9948851; PIDN:AAG06159.1; GSPDB:GN  
 A;Experimental source: strain PA01  
 A;Genetics:  
 A;Gene: PA2771

**RESULT 12**

H70922 hypothetical protein Rv3130c - *Mycobacterium tuberculosis* (strain H37RV)  
 C;Species: *Mycobacterium tuberculosis* revision  
 C;Accession: H70922  
 R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.  
 Nature 333, 537-544, 1998  
 A;Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
 A;Title: Deciphering the biology of *Mycobacterium tuberculosis* from the complete geno  
 A;Reference number: A70500; MUID:98295987; PMID:9634230  
 A;Accession: H70922  
 A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A;Molecule type: DNA  
 A;Residues: 1-463 <COL>  
 A;Cross-references: GB:Z95150; GB:AL123456; NID:93250708; PIDN:CAB08399.1; PID:e13010  
 A;Experimental source: strain H37RV  
 C;Genetics:  
 A;Gene: Rv3130c

**RESULT 13**

QY 95 PLQRURIV 102  
 Db 315 PLQRURIV 322

IMNG query Match 1.9%; Score 8; DB 2; length 463;  
 Best Local Similarity 100.0%; Pred. No. 13;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 R;Barthels, D.; Santoni, M.J.; Wille, W.; Ruppert, C.; Chaix, J.C.; Hirsch, M.R.; Fon  
 EMBO J. 6, 907-914, 1987  
 A;Title: Isolation and nucleotide sequence of mouse NCAM cDNA that codes for a Mr 79,  
 A;Reference number: A29673; MUID:87246524; PMID:3595563  
 A;Accession: A29673  
 A;Molecule type: mRNA  
 A;Residues: 1-725 <BAR>  
 A;Cross-references: EMBL:Y0051; NID:953342; PIDN:CAA68263.1; PID:953343  
 R;Barbas, J.A.; Chai, J.C.; Steinmetz, M.; Goridis, C.  
 EMBO J. 7, 625-632, 1988  
 A;Title: Differential splicing and alternative polyadenylation generates distinct NCA  
 A;Reference number: S00382; MUID:88283628; PMID:3396534  
 A;Accession: S00382

A; Molecule type: DNA  
A; Residues: 642-656, 'D', 658-725 <BA2>  
A; Cross-references: EMBL:X07195  
R; Rougon, G.; Marshak, D.R.  
J. Biol. Chem. 261: 3395-3401, 1986  
A; Title: Structural and immunological characterization of the amino-terminal domain of NCAM mediates cell-cell adhesion via homophilic binding with another NCAM molecule. Several forms of NCAM are produced by alternative splicing. See also PIR:JUMS  
C; Genetics:  
A; Gene: NCAM  
A; Map position: 9  
A; Introns: 7/0/1  
C; Superfamily: neural cell adhesion molecule; fibronectin type III repeat homology; immunoglobulin: signal sequence #status predicted <SIG>  
C; Keywords: alternative splicing; cell adhesion; duplication; heparin binding; membrane protein; immunoglobulin homology <IMM1>  
F; 1-19/Domain: immunoglobulin homology <IMM1>  
F; 132-191/Domain: immunoglobulin homology <IMM2>  
F; 161-195/Region: heparin binding #status predicted  
F; 228-290/Domain: immunoglobulin homology <IMM3>  
F; 263-272/Region: heparin binding #status predicted  
F; 228-290/Domain: immunoglobulin homology <IMM3>  
F; 263-272/Region: NCAM binding #status predicted  
F; 323-398/Domain: immunoglobulin homology <IMM4>  
F; 430-492/Domain: immunoglobulin homology <IMM4>  
F; 529-606/Domain: fibronectin type III repeat homology <FN3A>  
F; 635-695/Domain: fibronectin type III repeat homology <FN3B>  
F; 722-739/Domain: transmembrane #status predicted <TM>  
F; 740-858/Domain: intracellular #status predicted <INT>  
F; 411-96,139-189,235-288,330-396,437-490/Disulfide bonds: #status predicted  
F; 222,316,348,434,460,489/Binding site: carbohydrate (Asn) (covalent) #status predicted  
F; 41-96,139-189,235-288,330-386,427-480/Disulfide bonds: #status predicted  
F; 222,316,348,424,450,479/Binding site: carbohydrate (Asn) (covalent) #status predicted  
Query Match 1.9%; Score 8; DB 1; Length 725;  
Best Local Similarity 100.0%; Pred. No. 19;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 322 VTAEDGTO 329  
Db 98 VTAEDGTO 105

RESULT 14

IJRNc  
neuronal cell adhesion molecule short domain form precursor - rat  
N; Alternate names: NCAM-140  
C; Species: Rattus norvegicus (Norway rat)  
C; Date: 30-Sep-1991 #sequence\_revision 30-Sep-1991 #text\_change 22-Jun-1999  
C; Accession: S00846; B37795; I58136  
R; Small, S.J.; Shull, G.E.; Santoni, M.J.; Akeson, R.  
J. Cell Biol. 105, 2335-2445, 1987  
A; Title: Identification of a cDNA clone that contains the complete coding sequence for a  
A; Reference number: S00846; MUID:88059265; PMID:3680385  
A; Accession: S00846  
A; Molecule type: mRNA  
A; Residues: 1-859 <SMA>  
A; Cross-references: EMBL:X06564  
R; Small, S.J.; Akeson, R.  
J. Cell Biol. 111, 2089-2096, 1990  
A; Title: Expression of the unique NCAM VASE exon is independently regulated in distinct  
A; Reference number: A37795; MUID:91035620; PMID:1699951  
A; Accession: B37795  
A; Status: preliminary; not compared with conceptual translation  
A; Molecule type: mRNA  
A; Residues: 340-381 <SM2>  
R; Small, S.J.; Haines, S.L.; Akeson, R.A.  
Neuron 1, 1007-1017, 1988  
A; Title: Polypeptide variation in an N-CAM extracellular immunoglobulin-like fold is developmentally regulated in distinct  
A; Reference number: 158136; MUID:90166485; PMID:2483093  
A; Accession: I58136  
A; Status: preliminary; translated from GB/EMBL/DDJB  
A; Molecule type: DNA  
A; Residues: 355-364 <RES>  
A; Cross-references: GB:M32611; NID:9205643; PIDN:AAA41679.1; PID:9205644

RESULT 15

IJMSNL  
neuronal cell adhesion molecule 1 precursor, long domain splice form - mouse  
N; Alternate names: NCAM-180  
C; Contains: neural cell adhesion molecule, short domain splice form (NCAM-140)  
C; Species: Mus musculus (house mouse)  
C; Date: 31-Mar-1993 #sequence\_revision 31-Mar-1993 #text\_change 31-Dec-2000  
C; Accession: S00844; S00844; A20281; A44290; S00383  
R; Barthels, D.; Santoni, M.J.; Wille, W.; Ruppert, C.; Chaik, J.C.; Hirsch, M.R.; Fon  
EMBO J. 6, 907-914, 1987  
A; Title: Isolation and nucleotide sequence of mouse NCAM cDNA that codes for a Mr 79,000 protein  
A; Reference number: A29673; MUID:87246524; PMID:3595563  
A; Accession: A29673  
A; Molecule type: mRNA  
A; Residues: 1-548, 'T', 550-571, 'T', 573-574, 'D', 576-588, 'MQPSI', 593, 'S', 595-599, 'P', 601,  
A; Cross-references: EMBL:Y00051; MUID:g53342; PIDN:CAA60263.1; PMID:g53343  
R; Santoni, M.J.; Barthels, D.; Barbas, J.A.; Hirsch, M.R.; Steinmetz, M.; Goridis, C.  
Nucleic Acids Res. 15, 8621-8641, 1987  
A; Title: Analysis of cDNA clones that code for the transmembrane forms of the mouse neuronal cell adhesion molecule  
A; Reference number: S00844; MUID:88067687; PMID:3684567  
A; Accession: S00844  
A; Molecule type: mRNA  
A; Residues: 529-809, 1077-1115 <SAN>  
A; Cross-references: EMBL:X06328; NID:953322; PIDN:CAA29641.1; PID:9817984  
R; Barbas, J.A.; Chaix, J.C.; Steinmetz, M.; Goridis, C.  
EMBO J. 7, 625-632, 1988  
A; Title: Differential splicing and alternative polyadenylation generates distinct NCAM isoforms  
A; Reference number: S00382; MUID:88283628; PMID:3395534  
A; Accession: S00384  
A; Molecule type: DNA  
A; Residues: 642-1115 <BAR>  
A; Cross-references: EMBL:X07195  
R; Barthels, D.; Vopper, G.; Wille, W.  
Nucleic Acids Res. 16, 4217-4225, 1988  
A; Title: NCAM-180, the large isoform of the neural cell adhesion molecule of the mouse  
A; Reference number: A28281; MUID:88247737; PMID:2454455  
A; Accession: A28281  
A; Molecule type: mRNA  
A; Residues: 804-1081 <BA3>  
A; Cross-references: EMBL:X07244; NID:953321; PIDN:CAA30230.1; PID:9929720

R;Rougon, G.; Marshak, D.R.  
J. Biol. Chem. 261, 336-3401, 1986  
A;Title: structural and immunological characterization of the amino-terminal domain of n  
A;Reference number: A44290; MUID:86140120; PMID:3512556  
A;Molecule type: protein  
A;Residues: 20-36 <ROU>  
C;Comment: NCAM mediates cell-cell adhesion via homophilic binding with another NCAM mol  
C;Genetics:  
A;Gene: NCAM  
A;Map position: 9  
A;Introns: 6/3/3; 701/1; 770/2; 809/2; 1076/2  
C;Superfamily: neural cell adhesion molecule; fibronectin type III repeat homology; immu  
C;Keywords: alternative splicing; brain; cell adhesion; duplication; heparin binding; si  
F;1-19/Domain: signal sequence #status predicted <SIG>  
F;20-809/1077-1115/Product: neural cell adhesion molecule, long domain splice form #status experi  
F;20-711/Domain: extracellular #status predicted <EXT>  
F;114-98/Domain: immunoglobulin homology <IMM1>  
F;132-191/Domain: immunoglobulin homology <IMM2>  
F;152-156/Region: heparin binding #status predicted  
F;61-165/Region: heparin binding #status predicted  
F;228-290/Domain: immunoglobulin homology <IMM3>  
F;262-272/Region: NCAM binding #status predicted  
F;123-388/Domain: immunoglobulin homology <IMM4>  
F;420-482/Domain: immunoglobulin homology <IMM5>  
F;519-596/Domain: fibronectin type III repeat homology <FN3A>  
F;625-685/Domain: fibronectin type III repeat homology <FN3B>  
F;712-729/Domain: transmembrane #status predicted <TM>  
F;730-1115/Domain: intracellular #status Predicted <INT>  
F;41-96,139-89,235-286,330-386,427-480/Disulfide bonds: #status predicted  
F;222,316,348,424,450,479/Binding site: carbohydrate (Asn) (covalent) #status predicted  
  
Query Match  
Best Local Similarity 1.9%; Score 8; DB 1; Length 1115;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 322 VTAEDGTO 329  
Db 98 VTAEDGTO 105

Search completed: January 14, 2003, 17:15:20  
Job time : 45.3996 secs

**THIS PAGE BLANK (USPTO)**